Psychology, Public Policy, and Law
2005, Vol. 11, No. 2, 235–294
Copyright 2005 by the American Psychological Association
1076-8971/05/$12.00 DOI: 10.1037/1076-8971.11.2.235

THIRTY YEARS OF RESEARCH ON RACE DIFFERENCES IN COGNITIVE ABILITY

J. Philippe Rushton
The University of Western Ontario

Arthur R. Jensen
University of California, Berkeley

The *culture-only* (0% genetic–100% environmental) and the *hereditarian* (50% genetic–50% environmental) models of the causes of mean Black–White differences in cognitive ability are compared and contrasted across 10 categories of evidence: the worldwide distribution of test scores, *g* factor of mental ability, heritability, brain size and cognitive ability, transracial adoption, racial admixture, regression, related life-history traits, human origins research, and hypothesized environmental variables. The new evidence reviewed here points to some genetic component in Black–White differences in mean IQ. The implication for public policy is that the discrimination model (i.e., Black–White differences in socially valued outcomes will be equal barring discrimination) must be tempered by a distributional model (i.e., Black–White outcomes reflect underlying group characteristics).

Section 1: Background

Throughout the history of psychology, no question has been so persistent or so resistant to resolution as that of the relative roles of nature and nurture in causing individual and group differences in cognitive ability (Degler, 1991; Loehlin, Lindzey, & Spuhler, 1975). The scientific debate goes back to the mid-19th century (e.g., Galton, 1869; Nott & Glidden, 1854). Starting with the widespread use of standardized mental tests in World War I, average ethnic and racial group differences were found. Especially vexing has been the cause(s) of the 15-point Black–White IQ difference in the United States.

In 1969, the *Harvard Educational Review* published Arthur Jensen's lengthy article, "How Much Can We Boost IQ and School Achievement?" Jensen concluded that (a) IQ tests measure socially relevant general ability; (b) individual differences in IQ have a high heritability, at least for the White populations of the United States and Europe; (c) compensatory educational programs have proved generally ineffective in raising the IQs or school achievement of individuals or groups; (d) because social mobility is linked to ability, social class differences in IQ probably have an appreciable genetic component; and tentatively, but most controversially, (e) the mean Black–White group difference in IQ probably has some genetic component.

Jensen's (1969) article was covered in *Time*, *Newsweek*, *Life*, *U.S. News & World Report*, and *New York Times Magazine*. His conclusions, the theoretical issues they raised, and the public policy recommendations that many saw as stemming directly from them were dubbed "Jensenism," a term which entered the

J. Philippe Rushton, Department of Psychology, The University of Western Ontario, London, Ontario, Canada; Arthur R. Jensen, School of Education, University of California, Berkeley.

Correspondence concerning this article should be addressed to J. Philippe Rushton, Department of Psychology, The University of Western Ontario, London, Ontario N6A 5C2, Canada. E-mail: rushton@uwo.ca

dictionary. Since 1969, Jensen has continued to publish prolifically on all of these issues, and increasing numbers of psychometricians and behavioral geneticists have come to agree with one or more of the tenets of Jensenism (Snyderman & Rothman, 1987, 1988).

The Bell Curve (Herrnstein & Murray, 1994) presented general readers an update of the evidence for the hereditarian position along with several policy recommendations and an original analysis of 11,878 youths (including 3,022 Blacks) from the 12-year National Longitudinal Survey of Youth. It found that most 17-year-olds with high scores on the Armed Forces Qualification Test, regardless of ethnic background, went on to occupational success by their late 20s and early 30s, whereas those with low scores were more inclined to welfare dependency. The study also found that the average IQ for African Americans was lower than those for Latino, White, Asian, and Jewish Americans (85, 89, 103, 106, and 113, respectively; Herrnstein & Murray, 1994, pp. 273–278).

Currently, the 1.1 standard deviation difference in average IQ between Blacks and Whites in the United States is not in itself a matter of empirical dispute. A meta-analytic review by Roth, Bevier, Bobko, Switzer, and Tyler (2001) showed it also holds for college and university application tests such as the Scholastic Aptitude Test (SAT; $N = 2.4$ million) and the Graduate Record Examination (GRE; $N = 2.3$ million), as well as for tests for job applicants in corporate settings ($N = 0.5$ million) and in the military ($N = 0.4$ million). Because test scores are the best predictor of economic success in Western society (Schmidt & Hunter, 1998), these group differences have important societal outcomes (R. A. Gordon, 1997; Gottfredson, 1997).

The question that still remains is whether the cause of group differences in average IQ is purely social, economic, and cultural or whether genetic factors are also involved. Following publication of *The Bell Curve,* the American Psychological Association (APA) established an 11-person Task Force (Neisser et al., 1996) to evaluate the book's conclusions. Based on their review of twin and other kinship studies, the Task Force for the most part agreed with Jensen's (1969) *Harvard Educational Review* article and *The Bell Curve,* that within the White population the heritability of IQ is "around .75" (p. 85). As to the cause of the mean Black–White group difference, however, the Task Force concluded: "There is certainly no support for a genetic interpretation" (p. 97).

Among the factors contributing to the longstanding lack of resolution of this important and controversial issue are the difficulty of the subject matter, the political issues associated with it and the emotions they arouse, and the different meta-theoretical perspectives of the experimental and correlational methodologies. Cronbach (1957) referred to these conflicting approaches as the two "halves" of psychology because researchers are predisposed to draw different conclusions depending on whether they adopt a "manipulations-lead-to-change" or a "correlations-find-stability" paradigm.

Here we review in detail the research that has accumulated since Jensen's (1969) article and compare our findings with earlier reviews and evaluations such as those by Loehlin et al. (1975), P. E. Vernon (1979), Herrnstein and Murray (1994), the APA Task Force (Neisser et al., 1996), and Nisbett (1998). Facts in themselves typically do not answer scientific questions. For a question so complex as the cause of the average Black–White group difference in IQ, no one fact, one

study, nor indeed any single line of evidence, can hope to be determinative. Rather, resolving the issue requires examining several independent lines of evidence to determine if, when taken together, they confirm or refute rival hypotheses and research programs.

The philosophy of science methodology used here is guided by the view that, just as in individual studies the principal of aggregation holds *that a set of measurements provides a more reliable indicator than any single measure taken from the set* (Rushton, Brainerd, & Pressley, 1983), so in reviewing multiple lines of evidence, making strong inferences from a number of contending hypotheses is more efficacious than considering only one hypothesis at a time (Platt, 1964). Although strong inference is *the* method of science, it has, more often than not, been eschewed in this controversial debate.

The final section of this article addresses the question of what these conclusions imply for policy, specifically for the issues of educational and psychological testing, health, race relations, and conflicting worldviews about the essence of human nature. It suggests that the distributional model that takes genetic factors into account must temper the discrimination model that explains Black–White differences in socially valued outcomes.

Section 2: The Two Conflicting Research Programs

Here, we review the research on Black–White difference in average IQ published since Jensen's (1969) now 36-year-old article. We then apply the philosophy of science methodologies of Platt (1964), Lakatos (1970, 1978), and Urbach (1974a, 1974b) to determine if the preponderance of this new evidence strengthens or weakens Jensen's (1969) tentative assertion that it is more likely than not that some part of the cause of the mean Black–White difference is genetic. The data reviewed have been collated from articles in specialist journals and a number of scholarly monographs on the nature of intelligence, behavioral genetics, and social policy issues, as well as recent book-length reviews (Devlin, Feinberg, Resnick, & Roeder, 1997; Herrnstein & Murray, 1994; Jencks & Phillips, 1998; Jensen, 1998b; Lynn & Vanhanen, 2002; Rushton, 2000; Sternberg, 2000). While we focus on the mean Black–White difference in IQ because it is the topic on which most of the research to date has been conducted, studies of other traits (e.g., reaction times) and other groups (e.g., East Asians) are included when those data are sufficient and informative.

Some have argued that the cause of Black–White differences in IQ is a pseudo question because "race" and "IQ" are arbitrary social constructions (Tate & Audette, 2001). However, we believe these constructs are meaningful because the empirical findings documented in this article have been confirmed across cultures and methodologies for decades. The fuzziness of racial definitions does not negate their utility. To define terms, based on genetic analysis, roughly speaking, *Blacks* (Africans, Negroids) are those who have most of their ancestors from sub-Saharan Africa; *Whites* (Europeans, Caucasoids) have most of their ancestors from Europe; and *East Asians* (Orientals, Mongoloids) have most of their ancestors from Pacific Rim countries (Cavalli-Sforza, 2000; Cavalli-Sforza, Menozzi, & Piazza, 1994; Nei & Roychoudhury, 1993; Risch, Burchard, Ziv, & Tang, 2002). Although he eschewed the term *race*, Cavalli-Sforza's (2000, p. 70) maximum

likelihood tree made on the basis of molecular genetic markers substantially supports the traditional racial groups classification. Of course, in referring to population or racial group differences we are discussing *averages*. Individuals are individuals, and the three groups overlap substantially on almost all traits and measures.

The *hereditarian* position originated in the work of Charles Darwin (1859, 1871) and then was elaborated by his cousin Sir Francis Galton (1869, 1883). Based on research models used in behavioral genetics, this view contends that a substantial part (say 50%) of both individual and group differences in human behavioral traits is genetic. It therefore follows that even if all individuals in both groups were treated identically, average group differences would not disappear, though they might diminish.

The opposing *culture-only* position finds no need to posit any genetic causation, stating that if the environments for all individuals could be equalized, the observed group differences in average IQ would effectively disappear, though this might be difficult to achieve. This position has been predominant in the social sciences since the 1930s.

It is essential to keep in mind precisely what the two rival positions do and do not say—about a 50% genetic–50% environmental etiology for the hereditarian view versus an effectively 0% genetic–100% environmental etiology for the culture-only theory. *The defining difference is whether any significant part of the mean Black–White IQ difference is genetic rather than purely cultural or environmental in origin.* Hereditarians use the methods of quantitative genetics, and they can and do seek to identify the environmental components of observed group differences. Culture-only theorists are skeptical that genetic factors play any *independently effective* role in explaining group differences.

Most of those who have taken a strong position in the scientific debate about race and IQ have done so as either hereditarians or culture-only theorists. Intermediate positions (e.g., gene–environment interaction) can be operationally assigned to one or the other of the two positions depending on whether they predict any significant heritable component to the average group difference in IQ. For example, if gene–environment interactions make it impossible to disentangle causality and apportion variance, for pragmatic purposes that view is indistinguishable from the 100% culture-only program because it denies any potency to the genetic component proposed by hereditarians.

It is also important to define and interpret *heritability* correctly. Heritability refers to the genetic contribution to the *individual differences* (variance) in a particular group, not to the phenotype of a single individual. Heritability is not a constant that holds for all groups or in all environments. A heritability of 1.00 means all the observed differences in that group are due to genetic differences and not at all to their differences in the environment. A heritability of zero (0.00) means the converse. A heritability of 0.50 means the observed variation is equally the result of genetic and of environmental differences. The heritability of height in modern industrial populations, for example, is about 90%, which means that most of the differences in height among the individuals are due to their genetic differences.

Behavioral Genetics by Plomin, DeFries, McClearn, and McGuffin (2001) provides a detailed explanation of heritability (see also Jensen, 1973; Miele, 2002,

for general readers). Heritability estimates are true only for particular populations at particular times. They can vary in different populations or at different times. Equalizing environments, for example, produces the counterintuitive result of increasing heritability because any individual differences that remain must be due to genetic differences.

The cause of individual differences *within* groups has no necessary implication for the cause of the average difference *between* groups. A high heritability within one group does not mean that the average difference between it and another group is due to genetic differences, even if the heritability is high in both groups. However, within-groups evidence does imply the plausibility of the between-groups differences being due to the same factors, genetic or environmental. If variations in level of education or nutrition or genes reliably predict individual variation *within* Black and *within* White groups, then it would be reasonable to consider these variables to explain the differences *between* Blacks and Whites. Of course, independent evidence would then be needed to establish any relationship.

Heritability describes *what is* the genetic contribution to individual differences in a particular population at a particular time, not *what could be*. If either the genetic or the environmental influences change (e.g., due to migration, greater educational opportunity, better nutrition), then the relative impact of genes and environment will change. Heritability has nothing to say about *what should be*. If a trait has a high heritability it does not mean that it cannot be changed. Environmental change is possible. For example, phenylketunuria (PKU) is a single-gene disorder that causes mental retardation but that can be prevented by beginning a diet low in phenylaline early in life. (Note that the only effective treatment for PKU is aimed directly at the specific chemical factor that causes it.) The fact that the heritability of IQ is between 0.50 and 0.80 does not mean that individual differences are fixed and permanent. It *does* tell us that some individuals are genetically predisposed to be more teachable, more trainable, and more capable of changing than others, under current conditions (Jensen, 1973; Miele, 2002).

Having defined the terms of the debate, we now discuss approaches for resolving it. Lakatos's (1970, 1978) analytical methodology classifies research programs as being either progressive or degenerating. A *progressive* program not only explains existing phenomena and theoretical anomalies but also offers novel predictions, some of which can be tested and then either confirmed or rejected. A *degenerating* program merely accommodates existing anomalies by a series of new, unrelated, ad hoc hypotheses, ignores them, or denies their existence.

The philosopher Peter Urbach (1974a, 1974b) applied this methodology and concluded that the hereditarian/culture-only IQ debate is really a conflict of *research programs* that goes back to their classic proponents—Francis Galton (1869) for the hereditarians and J. B. Watson (1924) for the environmentalists. Each has an underlying set of assumptions, termed its *hard core*, and a *heuristic* machinery that generates hypotheses. The hard core of the hereditarian program is that (a) all individuals possess some level of general mental capacity called *general intelligence* that, to some degree, influences all cognitive activity, and (b) the differences between individuals and between groups in general intelligence are largely the result of genetic differences. The hard core of the culture-only program is that (a) there are a number of different learned mental skills or

intelligences, and (b) any observed differences in cognitive performance are the result of environmental factors.

Hereditarian heuristics include constructing better tests, developing better techniques for measuring mental abilities, and discovering biological correlates (e.g., heritability, inbreeding depression and heterosis, brain size, brain metabolic rate, brain evoked potentials, brain imaging) of these tests. The process then involves examining the similarities of the scores among people whose varying degrees of genetic resemblance can be predicted from Mendelian theory (Fisher, 1918). Culture-only heuristics include searching for the environmental factors that cause differences in intellectual performance and discovering the bias in existing tests. If two groups differ in mean IQ, culture-only theorists conjecture either that the lower scoring group has been exposed to one or more deleterious experience or been deprived of some beneficial environmental stimuli or that the tests are not valid measures of their true ability. Compensatory training might be initiated and the hypothesis confirmed if the groups then obtain more nearly equal scores, or if less biased tests are developed on which the group differences are reduced but still predict outside criteria. Of course, these two programs overlap to some degree, and a given experiment might well combine elements of the heuristics of each.

Reviewed here are new data sets for 10 categories of evidence that have become available since Jensen's (1969) article. They include the international pattern of IQ test scores, more and less *g*-loaded components of tests, heritability, brain-size and cognitive-ability relations, transracial adoption, racial admixture, regression to the mean, the race–behavior matrix, human origins research, and hypothesized environmental variables. These findings are then used to evaluate the culture-only and hereditarian models in terms of the methodology proposed by Lakatos (1970, 1978).

Section 3: Mean Race–IQ Differences: A Global Perspective

The IQ debate became worldwide in scope when it was shown that East Asians scored higher on IQ tests than did Whites, both within the United States and in Asia, even though IQ tests were developed for use in the Euro American culture (Lynn, 1977, 1978, 1982; P. E. Vernon, 1979, 1982). Around the world, the average IQ for East Asians centers around 106; that for Whites, about 100; and that for Blacks, about 85 in the United States and 70 in sub-Saharan Africa. Most of the early research was conducted in the United States, but some was also performed in Canada and the Caribbean (Eysenck, 1971, 1984; Jensen, 1969, 1973; Osborne & McGurk, 1982; Shuey, 1958, 1966; cf. Flynn, 1980; Kamin, 1974; Lewontin, Rose, & Kamin, 1984). In the United States, 15% to 20% of the Black IQ distribution exceeds the White median IQ, so many Blacks obtain scores above the White average. This same order of mean group differences is also found on "culture-fair" tests and on reaction time tasks. Hundreds of studies on millions of people have confirmed the *three-way* racial pattern (Jensen, 1998b; Lynn & Vanhanen, 2002; Rushton, 2000).

Racial-group differences in IQ appear early. For example, the Black and the White 3-year-old children in the standardization sample of the Stanford–Binet IV show a 1 standard deviation mean difference after being matched on gender, birth order, and maternal education (Peoples, Fagan, & Drotar, 1995). Similarly, the

Black and the White 2½- to 6-year-old children in the U.S. standardization sample of the Differential Aptitude Scale have a 1 standard deviation mean difference. No data are available for East Asian children at the youngest ages. On the Differential Aptitude Battery, by age 6, however, the average IQ of East Asian children is 107, compared with 103 for White children and 89 for Black children (Lynn, 1996). The size of the average Black–White difference does not change significantly over the developmental period from 3 years of age and beyond (see Jensen, 1974, 1998b).

Serious questions have been raised about the validity of using tests for racial comparisons. However, because the tests show similar patterns of internal item consistency and predictive validity for all groups, and because the same differences are found on relatively culture-free tests, many psychometricians have concluded that the tests *are* valid measures of racial differences, at least among people sharing the culture of the authors of the test (Jensen, 1980; Wigdor & Garner, 1982). This conclusion was endorsed by the APA Task Force's statement: "Considered as predictors of future performance, the tests do not seem to be biased against African Americans" (Neisser et al., 1996, p. 93).

Most disputed is the validity of the low mean IQ scores reported for sub-Saharan Africans. Lynn's (1991) review of 11 studies found a mean IQ of 70. A subsequent review of over two dozen studies by Lynn and Vanhanen (2002) found an average IQ of 70 for West, Central, East, and Southern Africa. For example, in Nigeria, Fahrmeier (1975) collected data on 375 children ages 6 to 13 years in a study of the effects of schooling on cognitive development. The children's mean score on the Colored Progressive Matrices was 12 out of 36, which is at the 4th percentile for 9½-year-olds on U.S. norms, or an IQ equivalent of about 75 (Raven et al., 1990, pp. 97–98). In Ghana, Glewwe and Jacoby (1992) reported a World Bank study that tested a representative sample of 1,736 individuals ranging in age from 11 to 20 years old from the entire country. All had completed primary school; half were attending middle school. Their mean score on the Colored Progressive Matrices was 19 out of 36, which is below the 1st percentile for 15½-year-olds on U.S. norms, an IQ equivalent of less than 70. In Kenya, Sternberg et al. (2001) administered the Colored Progressive Matrices to 85 children ages 12 to 15 years old who scored 23.5 out of 36, which is about the 2nd percentile for 13½-year-olds, an IQ equivalent of 70. In Zimbabwe, Zindi (1994) reported mean IQs for 204 African 12- to 14-year-olds of 67 on the Wechsler Intelligence Scale for Children—Revised (WISC–R) and of 72 on the Standard Progressive Matrices. In South Africa, Owen (1992) found that 1,093 African 12- to 14-year-old high school students solved 28 out of 60 problems on the Standard Progressive Matrices, which is around the 10th percentile, or an IQ equivalent of about 80 (Raven, Raven, & Court, 1998, p. 77). Again in South Africa, Skuy, Schutte, Fridjhon, and O'Carroll (2001) found mean scores 1 to 2 standard deviations below U.S. norms on a wide variety of individually administered tests given to 154 African high school students under optimized conditions.

Black university students in South Africa also show relatively low mean test scores. Sixty-three undergraduates at the all-Black universities of Fort Hare, Zululand, the North, and the Medical University of South Africa had a full-scale IQ of 77 on the Wechsler Adult Intelligence Scale—Revised (Avenant, 1988, cited in Nell, 2000, pp. 26–28). In a study at the University of Venda in South

Africa's Northern Province by Grieve and Viljoen (2000), 30 students in 4th-year law and commerce averaged a score of 37 out of 60 on the Standard Progressive Matrices, equivalent to an IQ of 78 on U.S. norms. A study at South Africa's University of the North by Zaaiman, van der Flier, and Thijs (2001) found the highest scoring African sample to that date—147 first-year mathematics and science students who scored 52 out of 60 on the Standard Progressive Matrices, which is equivalent to an IQ of 100. This higher score may reflect the fact that they were mathematics and science students, specially selected for admission to the university from a pool of 700 applicants on the basis of a math-science selection test.

At the University of the Witwatersrand in Johannesburg, South Africa, Rushton, Skuy, and colleagues gave the Raven's Progressive Matrices in four separate studies under optimal testing conditions. Rushton and Skuy (2000) found 173 African 1st-year psychology students averaged an IQ equivalent of 84. Skuy et al. (2002) tested another 70 psychology students who averaged an IQ equivalent of 83. After receiving training on how to solve Matrices-type items, their mean score rose to an IQ equivalent of 96. Rushton, Skuy, and Fridjhon (2002, 2003) gave nearly 200 African 1st-year engineering students both the Standard and the Advanced version of the Raven's test and found they averaged an IQ of 97 on the Standard and 103 on the Advanced, making them the highest scoring African sample on record. (The White university students in these four studies had IQs from 105 to 117; East Indian students had intermediate IQs, from 102 to 106.)

Many critics claim that Western-developed IQ tests are not valid for groups as culturally different as sub-Saharan Africans (e.g., Nell, 2000). The main evidence to support a claim of *external* bias would be if the test failed to predict performance for Africans. Even if tests only *underpredicted* performance for Africans compared with non-Africans, it would suggest that their test scores underestimated their "true" IQ scores. However, a review by Kendall, Verster, and von Mollendorf (1988) showed that test scores for Africans have about equal predictive validity as those for non-Africans (e.g., 0.20 to 0.50 for students' school grades and for employees' job performance). The review also showed that many of the factors that influence scores in Africans are the same as those for Whites (e.g., coming from an urban vs. a rural environment; being a science rather than an arts student; having had practice on the tests; and the well-documented curvilinear relationship with age). Similarly, Sternberg et al.'s (2001) study of Kenyan 12- to 15-year-olds found that IQ scores predicted school grades, with a mean $r = .40$ ($p < .001$; after controlling for age and socioeconomic status [SES], $r = .28$, $p < .01$). In Rushton et al.'s (2003) study of African and non-African engineering students at the University of the Witwatersrand, scores on the Advanced Progressive Matrices correlated with scores on the Standard Progressive Matrices measured 3 months earlier (.60 for Africans; .70 for non-Africans) and with end-of-year exam marks measured 3 months later (.34 for Africans; .28 for non-Africans). Figure 1 shows the regression of exam marks on test scores for these university students.

Although predictive validity is the ultimate pragmatic criterion for absence of bias, critics also suggest that the items have different meanings for Africans than they do for Whites or East Indians (Nell, 2000). This hypothesis of *internal* bias has been tested. The psychometric studies by Owen (1992) on thousands of high

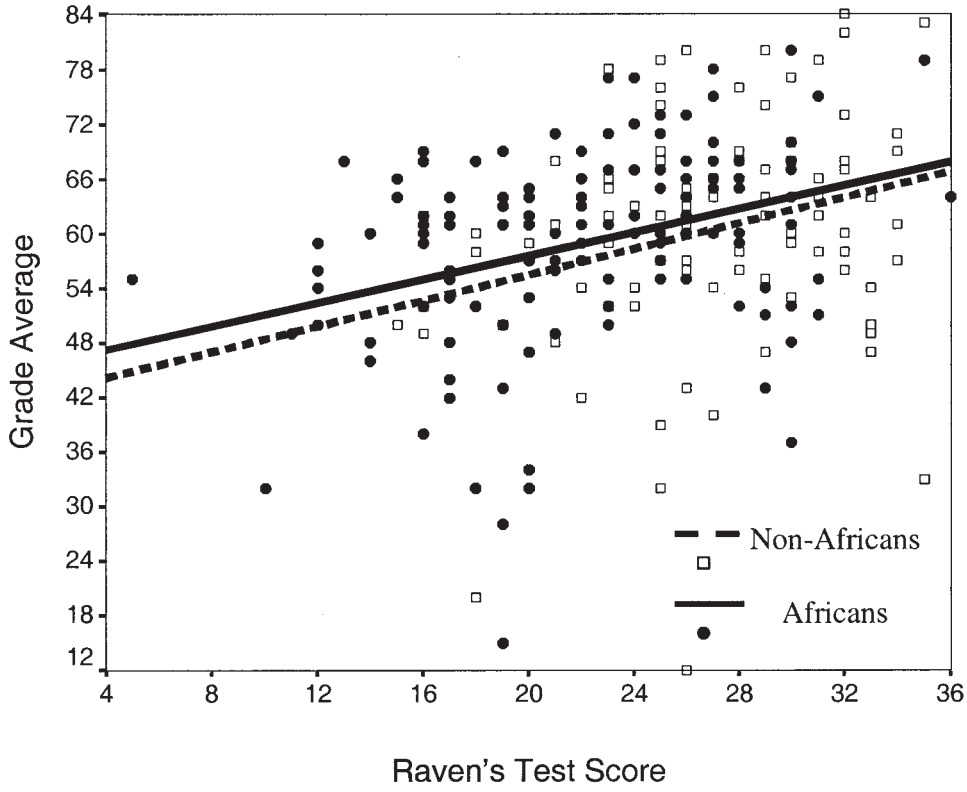

Figure 1. Regression of Raven's scores on university grades for Africans and non-Africans. From "Performance on Raven's Advanced Progressive Matrices by African, East Indian, and White Engineering Students," by J. P. Rushton, M. Skuy, and P. Fridjhon, 2003, *Intelligence, 31,* p. 133. Copyright 2003 by Elsevier Science. Reprinted with permission.

school students, and by Rushton and Skuy (2000; Rushton et al., 2002, 2003) on hundreds of university students, found almost identical item structures in Africans, Whites, and East Indians on the Progressive Matrices. Items found difficult by one group were difficult for the others; items found easy by one group were easy for the others (mean rs $= .90$, $p < .001$). The item–total score correlations for Africans, Whites, and East Indians were also similar, indicating that the items measured similar psychometric constructs in all three groups. (Section 4 reviews evidence of the similarity of the *g* factor in Africans and non-Africans.) The only reliable example of bias so far discovered in this extensive literature is the rather obvious internal bias on the Vocabulary components of tests such as the Wechsler for groups that do not have English as their first language (e.g., Skuy et al., 2001). Even here, the language factor only accounts for about 0.5 of a standard deviation, out of the overall 2.0 standard deviation difference, between Africans and Whites.

Could it make a difference that Africans have less experience in solving problems such as those on the Raven's, are less testwise, and have less access to coaching than non-Africans? Raven (2000) showed that students who were

encouraged to engage in complex cognitive tasks increased in self-direction, understanding, and competence. In South Africa, Skuy and Shmukler (1987) applied Feuerstein's (1980) Mediated Learning Experience and raised the Raven scores of Black high school students. Skuy, Hoffenberg, Visser, and Fridjhon (1990) found generalized improvements for Africans with what they termed a *facilitative temperament*. In an intervention study with 1st-year psychology students at the University of the Witwatersrand, Skuy et al. (2002) increased Raven's test scores in both Africans and non-Africans after intervention training. Both experimental groups improved over the baseline compared with their respective control groups, with significantly greater improvement for the African group (IQ score gains of 83 to 97 in Africans; 103 to 107 in non-Africans). The question remains, however, whether such intervention procedures only increase performance through mastery of subject-specific knowledge or whether they increase *g*-like problem-solving ability that generalizes to other tests as well (te Nijenhuis, Voskuijl, & Schijve, 2001).

Some argue that African students are less interested, more anxious, work less efficiently, or give up sooner on items they find difficult, perhaps because the problems have less meaning for them (e.g., Nell, 2000). Four findings argue against these hypotheses. First, Rushton and Skuy (2000) closely observed the test-taking behavior of Africans and noted that they worked very diligently, typically staying longer than Whites to recheck their answers. Second and third, there are the similar predictive validities and internal consistencies previously discussed. Finally, there is supporting evidence from reaction-time research.

Reaction time is one of the simplest culture-free cognitive measures. Most reaction time tasks are so easy that 9- to 12-year-old children can perform them in less than 1 s. But even on these very simple tests, children with higher IQ scores perform faster than do children with lower scores, perhaps because reaction time measures the neurophysiological efficiency of the brain's capacity to process information accurately—the same ability measured by intelligence tests (Deary, 2000; Jensen, 1998b). Children are not trained to perform well on reaction time tasks (as they are on certain paper-and-pencil tests), so the advantage of those with higher IQ scores on these tasks cannot arise from practice, familiarity, education, or training.

For three reaction time tasks (the simple, choice, and odd-man-out tasks), individuals with higher IQ scores average faster and less variable reaction times. For any one task, the correlation between reaction time and IQ normally lies between .20 and .40. A review of several studies concluded that the six measures combined (i.e., the average time and the variability for the three reaction time tasks) produce a multiple correlation of .67 (Deary, 2000). This is about the same magnitude as the correlation between two conventional intelligence tests of, say, reasoning ability and vocabulary.

Lynn and his colleagues carried out a series of reaction time studies on over 1,000 nine-year-old East Asian children in Japan and Hong Kong, White children in Britain and Ireland, and Black children in South Africa (summarized by Lynn & Vanhannen, 2002, pp. 66–67). The Progressive Matrices were given as a nonverbal test of intelligence, along with the simple, choice, and odd-man-out tasks. Reaction times and variabilities were measured by computer and hence were not subject to any human error in recording. For details, see Shigehisa and

Lynn (1991) for Japan; Chan and Lynn (1989) for Hong Kong and Britain; Lynn (1991) for Ireland; and Lynn and Holmshaw (1990) for South Africa.

The correlations between IQ and reaction times for the five countries are summarized in Table 1. The East Asian children in Hong Kong and Japan obtained the highest IQs, followed in descending order by the White children in Britain and Ireland, and then the Black children in South Africa. The medians for simple reaction time, choice reaction time, and odd-man-out reaction time follow the same descending order as the IQs. Because all the tasks take less than 1 s, all children found them easy. The variabilities in the three reaction time measures for the three groups follow the same general descending trend.

The same pattern of average scores on these and other reaction time tasks (i.e., East Asians faster than Whites faster than Blacks) is found within the United States. Jensen (1993) and Jensen and Whang (1994) examined the time taken by over 400 schoolchildren ages 9 to 12 years old in California to retrieve overlearned addition, subtraction, or multiplication of single digit numbers (from 1 to 9) from long-term memory. All of the children had perfect scores on paper-and-pencil tests of this knowledge, which was then reassessed using the Math Verification Test. The response times significantly correlated (negatively) with Raven Matrices scores, whereas movement times have a near-zero correlation. The average reaction times for the three racial groups differ significantly (see Figure 2). They cannot be explained by the groups' differences in motivation because the East Asian children averaged a *shorter* response time but a *longer* movement time than did the Black children.

Section 4: The *g* Factor and Mean Race–IQ Differences

Jensen (1998b) showed that a test's *g* loading (*g* being the general factor of intelligence) is the best predictor, not just of that test's correlation with scholastic and workplace performance, but of heritability coefficients determined from twin

Table 1
Sample Size, Mean IQ Score, and Reaction Time Measures (in Milliseconds) From Five Countries, and the Reaction Time Correlations with IQ

Variable	Hong Kong	Japan	Britain	Ireland	South Africa	*SD*	*r*[a]
Sample size	118	110	239	317	350	—	—
IQ scores	113	110	100	89	67	—	—
Simple reaction time	361	348	371	388	398	64	.94*
Choice reaction time	423	433	480	485	489	67	.89*
Odd-man-out reaction time	787	818	898	902	924	187	.96*
Variability of simple reaction time	99	103	90	121	139	32	.83*
Variability of choice reaction time	114	138	110	141	155	30	.73*
Variability of odd-man-out reaction time	269	298	282	328	332	95	.85*

Note. Data in this table are reconfigured from *IQ and the Wealth of Nations* (p. 67, Table 6.2), by R. Lynn and T. Vanhanen, 2002, Westport, CT: Praeger. Copyright 2002 by R. Lynn and T. Vanahanen. Adapted with permission.
[a] r = reliability.
* $p < .05$.

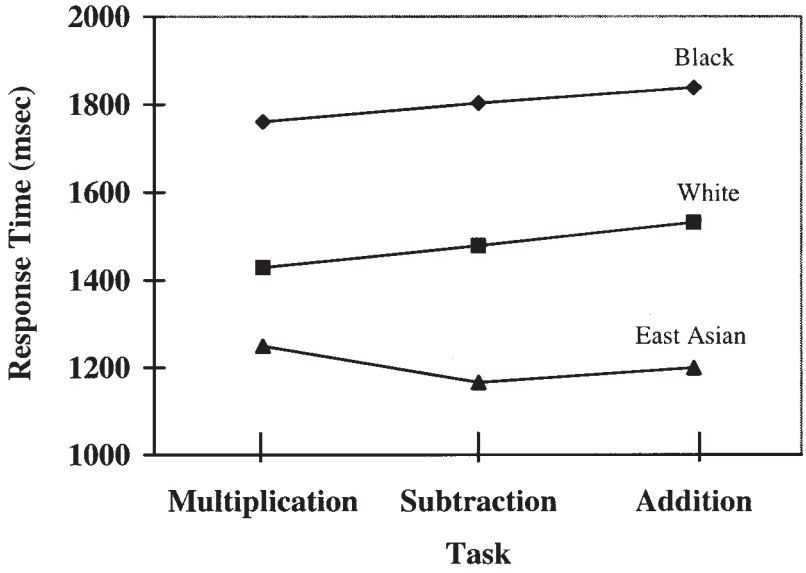

Figure 2. Mean response times of 10-year-old Black, White, and East Asian children on the Math Verification Test for Multiplication, Subtraction, and Addition. Redrawn by A. R. Jensen from data in Jensen (1993; Jensen & Whang, 1994). Copyright 2002 by A. R. Jensen. Reprinted with permission.

studies, inbreeding depression scores calculated in children of cousin-marriages, brain evoked potentials, brain pH levels, brain glucose metabolism, as well as nerve conduction velocity, reaction time, and other physiological factors. These correlations argue strongly for the heritable and biological, as opposed to the mere statistical reality of *g*. Because the mean Black–White group difference in IQ is more pronounced on high-*g*-loaded tests than it is on low-*g*-loaded tests, it suggests that it is not the result of any idiosyncratic cultural peculiarities of this or that test.

Spearman (1927, p. 379) first proposed the hypothesis that the mean Black–White group difference in IQ would be "most marked in just those [tests] which are known to be saturated with *g*." Jensen (1980, p. 535) designated it as "Spearman's hypothesis" and developed the *method of correlated vectors* to test it. This method correlates the standardized Black–White mean differences on a set of cognitive tests with their respective *g* loadings, a significant positive correlation supporting the hypothesis. The rationale is straightforward. If *g* is the main source of between- *and* within-group differences, then there should be a positive relationship between a given test's *g* loading and the mean Black–White group difference on that test: The more *g*-loaded the test, the greater the Black–White group difference on that test. A corollary is the prediction that when race (scored as Blacks = 1, Whites = 2) is factor analyzed along with scores from a number of diverse cognitive tests, its highest loading on the resulting correlation matrix will be with the *g* factor.

Jensen (1998b, pp. 369–379) summarized 17 independent data sets of nearly 45,000 Blacks and 245,000 Whites derived from 149 psychometric tests and found that the *g* loadings consistently predicted the magnitude of the mean

Black–White group difference ($r = .62$, $p < .05$). This was borne out even among 3-year-olds administered eight subtests of the Stanford–Binet in which the rank correlation between *g* loadings and the mean Black–White group differences was .71 ($p < .05$; Peoples et al., 1995). Subsequently, Nyborg and Jensen (2000) analyzed a unique battery of 19 highly diverse cognitive tests from an archival data set of 4,462 males who had served in the U.S. Armed Forces. The *g* factor was extracted using different methods. Spearman's hypothesis was confirmed, with an average correlation of .81 between the race difference on a test and its *g* loading. Nyborg and Jensen concluded that Spearman's original conjecture about the mean Black–White difference on the *g* factor "should no longer be regarded as just an hypothesis but as an empirically established fact" (p. 599). Only one study to date has examined East Asian–White difference on psychometric tests as a function of their *g* loadings; it confirmed the hypothesis for 15 cognitive tests administered to two generations of Americans of Japanese, Chinese, and European ancestry. In this case, the more *g*-loaded the test, the greater the mean East Asian–White group difference favoring East Asians (Nagoshi, Johnson, DeFries, Wilson, & Vandenberg, 1984).

Studies in Southern Africa have also found the mean Black–White IQ difference is mainly on *g*. Lynn and Owen (1994) were the first to test explicitly Spearman's hypothesis in sub-Saharan Africa, administering the Junior Aptitude Test to 1,056 White, 1,063 Indian, and 1,093 Black 16-year-old high school students in South Africa. They found a 2 standard deviation difference between the Africans and Whites (yielding an average African IQ of about 70) and a 1 standard deviation difference between the Whites and Indians (yielding an average Indian IQ of 85). They then tested Spearman's hypothesis and found the African–White differences correlated .62 ($p < .05$) with the *g* factor extracted from the African sample, but only .23 with *g* extracted from the White sample. They did not find any White–Indian differences on *g*.

Jensen (1998b, p. 388) noted some problems with Lynn and Owen's (1994) South African study, but their results on Black–White differences have been well corroborated since then and extended to include East Indians and "Coloreds" (the term used to refer to the mixed-race population of South Africa). Thus, Rushton (2001) reanalyzed data on 10 subtests of the WISC–R published on 154 high school students in South Africa by Skuy et al. (2001) and found African–White differences were mainly on *g*. Rushton and Jensen (2003) compared data on the WISC–R from 204 African 12- to 14-year-olds from Zimbabwe published by Zindi (1994) with the U.S. normative sample for Whites and found 77% of the between-groups race variance was attributable to a single source, namely *g*.

Spearman's hypothesis has been confirmed in South Africa using test item analyses as well. Rushton and Skuy (2000) studied 309 university students at the University of the Witwatersrand and found that the more an individual item from the Raven's Standard Progressive Matrices measured *g* (estimated by its item–total correlation), the more it correlated with the standardized African–White difference on that item. Rushton (2002) analyzed the item data from 4,000 high school students in South Africa on Raven's Standard Progressive Matrices published by Owen (1992) and found the four-way African–Colored–East Indian–White differences were all on *g*. In two studies of engineering students, Rushton et al. (2002, 2003) found that the more the items from both the Standard and the

Advanced Progressive Matrices loaded on *g*, the better they predicted the magnitude of African–East Indian–White differences. The *g* loadings showed cross-cultural generality; those calculated on the East Indian students predicted the magnitude of the African–White differences.

Spearman's hypothesis was also confirmed when the *g* factor was extracted from 12 reaction time variables given to the 820 nine- to twelve-year-olds. While all of the children could do the tasks in less than 1 s, the correlations between the *g* loadings and the mean Black–White differences on the reaction time tasks range from .70 to .81 (Jensen, 1993). These results bear out Spearman's hypothesis even more strongly than do those from conventional psychometric tests. The hypothesis that the mean Black–White group difference on these tests reflects a difference in motivation is again disconfirmed by the fact that although Whites averaged faster *reaction* times than Blacks, Blacks averaged faster *movement* times than Whites. And again, East Asians typically averaged higher than Whites on the *g* factor extracted from their (faster) reaction time measures (Jensen & Whang, 1994).

Spearman's hypothesis, as demonstrated by the method of correlated vectors, cannot be a chimera or a methodological artifact, as a few critics have claimed (e.g., Gould, 1996, p. 350; Schönemann, 1992). In the method of correlated vectors, the means and standard deviations of the variables cannot have any mathematical relationship with the factor structure of the correlation matrix because the means and the variances of all the tests in the factor-analyzed correlation matrix are totally removed by the Pearson correlations, which convert all variables to *z* scores. Therefore, any systematic relationship between factor loadings and standardized group means (or group mean differences) must be an empirical fact, not an artifact (Jensen, 1992).

Other claims of artifact are contradicted by Dolan's (1997) technical commentaries on the method of correlated vectors (Dolan, 1997, 2000). Dolan argued that the method of correlated vectors is not incorrect but that it lacks *specificity*; that is, it does not incorporate tests of alternative models of the factor structure of group differences or incorporate statistical tests to compare them for goodness-of-fit. In its place, he advocated use of the multigroup confirmatory factor model for testing Spearman's hypothesis. Statistical tests of significance are a built-in feature of this procedure. Dolan and Hamaker (2001) have applied it to two large published data sets (Jensen & Reynolds, 1982; Naglieri & Jensen, 1987). The results statistically confirmed the conclusion derived from the method of correlated vectors regarding a "weak form" of Spearman's hypothesis: Black–White group differences were predominantly on the *g* factor, although the groups also showed differences on some lower order factors (e.g., short-term memory and spatial ability) independent of *g*.

Section 5: Gene–Environment Architecture and Mean Black–White IQ Differences

Dozens of twin, adoption, and family studies have confirmed the high heritability of intellectual and behavioral traits, and even reaction time tasks, within a race (Bouchard, 1996; Bouchard & Loehlin, 2001; Deary, 2000; Plomin et al., 2001). Most of these estimates have been calculated on White samples. One study of 543 pairs of identical and 134 pairs of nonidentical 12-year-old Japanese twins

in Japan reported a substantial heritability of 0.58 for IQ (Lynn & Hattori, 1990). The hereditarian model views race differences simply as aggregated individual differences of this sort.

The culture-only model, however, predicts that special factors such as poverty, the history of slavery, and White racism have operated on the Black population and suppressed natural levels of intelligence and so made heritabilities in Blacks substantially lower than they are in Whites. In arguing against Galton's (1869) hereditarian position, Charles H. Cooley (1897), a founder and first president of the American Sociological Association, was the first to introduce the powerful analogy that corn seeds given a normal environment grow plants of full height whereas seeds given a deprived environment grow plants of stunted height. According to this view, cultural *deprivation*, not heredity, is the cause of any Black–White IQ differences.

It is an empirical question whether heritabilities are the same for Blacks as for Whites. Loehlin et al. (1975, pp. 114–116) reviewed the literature to that date and found that while there was some evidence suggesting a lower heritability of intelligence for Blacks than for Whites (e.g., Scarr-Salapatek, 1971), a larger body of evidence suggested equal heritabilities in the two groups. Subsequently, Osborne's (1980) Georgia Twin Study compared 123 Black and 304 White pairs of 12- to 18-year-old twins drawn from schools in Georgia, Kentucky, and Indiana, given the Basic Test Battery, along with smaller subsets of twins given the Primary Mental Abilities test and the Cattell Culture Fair Intelligence test. Osborne found heritabilities of about 50% for both Blacks and Whites, all significantly different from zero but not from each other. (The heritabilities of the Basic, Primary, and Cattell tests were, respectively, for Whites, 0.61, 0.37, and 0.71, and for Blacks, 0.75, 0.42, and 0.19; Osborne, 1980, pp. 68–69, 89, 98.)

Another way of answering the question is to compare their psychometric factor structures of kinship patterns, background variables, and subtest correlations. If there are minority-specific developmental processes arising from cultural background differences between the races at work, they should be reflected in the correlations between the background variables and the outcome measures. Rowe (1994; Rowe, Vazsonyi, & Flannery, 1994, 1995) examined this hypothesis in a series of studies using structural equation models. One study of six data sources compared cross-sectional correlational matrices (about 10×10) for a total of 8,528 Whites, 3,392 Blacks, 1,766 Hispanics, and 906 Asians (Rowe et al., 1994). These matrices contained both independent variables (e.g., home environment, peer characteristics) and developmental outcomes (e.g., achievement, delinquency). A LISREL goodness-of-fit test found each ethnic group's covariance matrix equal to the matrix of the other groups. Not only were the Black and White matrices nearly identical, but they were as alike as the covariance matrices computed from random halves within either group. There were no distortions in the correlations between the background variables and the outcome measures that suggested any minority-specific developmental factor.

Another study examined longitudinal data on academic achievement (Rowe et al., 1995). Again, any minority-specific cultural processes affecting achievement should have produced different covariance structures among ethnic and racial groups. Correlations were computed between academic achievement and family environment measures in 565 full-sibling pairs from the National Longitudinal

Survey of Youth, each tested at ages 6.6 and 9.0 years (White $N = 296$ pairs; Black $N = 149$ pairs; Hispanic $N = 120$ pairs). Each racial group was treated separately, yielding three 8×8 correlation matrices, which included age as a variable. Because LISREL analysis showed the matrices were equal across the three groups, there was no evidence of any special minority-specific developmental process affecting either base rates in academic achievement or any changes therein over time.

A nearly identical statistical structure on intellectual variables across ethnic and racial groups has been reported in large-scale studies of military samples. Ree and Carretta (1995) examined a nationally representative sample of young Black, White, and Hispanic men and women who took the Armed Services Vocational Aptitude Battery (ASVAB; $N = 9{,}173$). The ASVAB, which is used to select applicants for all military enlistments and assign them to first jobs, consists of 10 separately scored subtests (General Science, Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Numerical Operations, Coding Speed, Auto and Shop Information, Mathematics Knowledge, Mechanical Comprehension, and Electronics Information). Despite the especially wide variety of subtests, Ree and Carretta found the hierarchical factor structure of ASVAB subtest scores was virtually identical across the three groups. Similarly, Carretta and Ree (1995) examined the more specialized and diverse Air Force Officer Qualifying Test, a multiple-aptitude battery that had been given to 269,968 applicants (212,238 Whites, 32,798 Blacks, 12,647 Hispanics, 9,460 Asian Americans, and 2,551 Native Americans). The *g* factor accounted for the greatest amount of variance in all groups, and its loadings differed little by ethnicity. Thus, the factor structure of cognitive ability is nearly identical for Blacks and for Whites, as was found in the studies by Owen (1992) and Rushton and Skuy (2000; Rushton et al., 2002, 2003) comparing Africans, East Indians, and Whites on the item structures of tests described in Section 3. There was no "Factor X" specific to race.

Within-race heritabilities have also been calculated using structural equation modeling. Rowe and Cleveland (1996) estimated the genetic architecture for Black and White full- and half-siblings from the National Longitudinal Survey of Youth (106 pairs of Black half-sibs, 53 pairs of White half-sibs; 161 pairs of Black full-sibs, 314 pairs of White full-sibs). Three Peabody Individual Achievement Tests were used (Mathematics, Reading Comprehension, and Reading Recognition). The best-fitting model was one in which the sources of the differences between individuals within race and of the differences between racial means was the same—about 50% genetic and 50% environmental. Similarly, Jensen (1998b, p. 465) used structural equation modeling to reanalyze a subset of the Georgia Twin Study (comprising 123 Black and 304 White pairs of 12- to 18-year-old twins). He broke down the phenotypic mean differences into their genetic and environmental contributions and tested four alternative models: *only* genetic factors, *only* environmental factors, *neither* genes *nor* environment, and genes *plus* environment. The model of both genetic *and* environmental factors best explained the observed Black–White group differences in IQ, whereas both the genetic-only and the environmental-only explanations were inadequate.

Heritability data are especially informative when the hereditarian and the culture-only models make opposite predictions. For example, the hereditarian model predicts race differences will be greater on those subtests that are more

heritable within races, whereas culture-only theory predicts they will be greater on subtests that are more culturally malleable (i.e., those with lower heritabilities) on which races should grow apart as a result of dissimilar experiences. Analyses of several independent data sets support the genetic hypothesis.

Nichols (1972, cited in Jensen, 1973, pp. 116–117) was the first to apply differential heritabilities in the study of racial-group differences. He estimated the heritability of 13 tests from 543 pairs of 7-year-old siblings, including an equal number of Blacks and Whites, and found a .67 correlation between the heritability of a test and the magnitude of the Black–White group difference on that test. Subsequently, Jensen (1973, pp. 103–119) calculated the *environmentality* of a test (defined as the degree to which sibling correlations departed from the pure genetic expectation of 0.50) in Black and in White children and found it was inversely related to the magnitude of the Black–White group difference (r = –.70); that is, the more environmentally influenced a test, the less pronounced its Black–White group difference.

Prompted by Jensen's approach, Rushton (1989) estimated genetic influence from the amount of inbreeding depression found on the 11 tests of the WISC. Inbreeding depression occurs in offspring who receive the same harmful recessive genes from each of their closely related parents. Rushton found a positive correlation between inbreeding depression scores calculated from 1,854 cousin-marriages in Japan and the magnitude of the mean Black–White group difference in the United States on the same 11 Wechsler tests (.48). This contradicts culture-only theory, which predicts that mean differences between Blacks and Whites should be greater on those subtests most affected by the environment (i.e., those showing the lowest amount of inbreeding depression). We know of no nongenetic explanation for the relation between inbreeding depression scores from Japan and mean Black–White group differences in the United States.

Other aspects of the gene–environment architectural matrix also pertain to the question of mean Black–White group differences. First, it is possible to distinguish between two different types of environmental effects. S*hared* (also called *common* or *between-family*) environmental effects are due to variables all children reared in the same family (whether genetically related or adopted) have in common but that differ between families (e.g., father's occupation, family cultural practice, parents' child-rearing style). *Nonshared* (also called *unique* or *within-family*) effects are specific to each child in the same family and therefore differ within families (e.g., an accident, illness, or chance friendship that happens to one sibling and not to the other). Twin and adoption studies can be used to measure the two types of environmental effect (Plomin, DeFries, & Loehlin, 1977; see also Plomin & Daniels, 1987; Plomin et al., 2001).

Based on within-race data, Figure 3 summarizes the changes with age in the proportions of the total IQ variance attributable to genetic factors and to the effects of the shared and the nonshared environment. It is based on an analysis of 6,370 monozygotic and 7,212 dizygotic twin pairs reared together (McGue, Bouchard, Iacona, & Lykken, 1993). As can be seen, the estimated proportion of IQ variance associated with shared environmental factors is relatively constant at approximately 30% for ages up to 20 years but then drops to 0% in adulthood. The estimated proportion of IQ variance associated with genetic factors increases throughout development, but especially after 20 years of age.

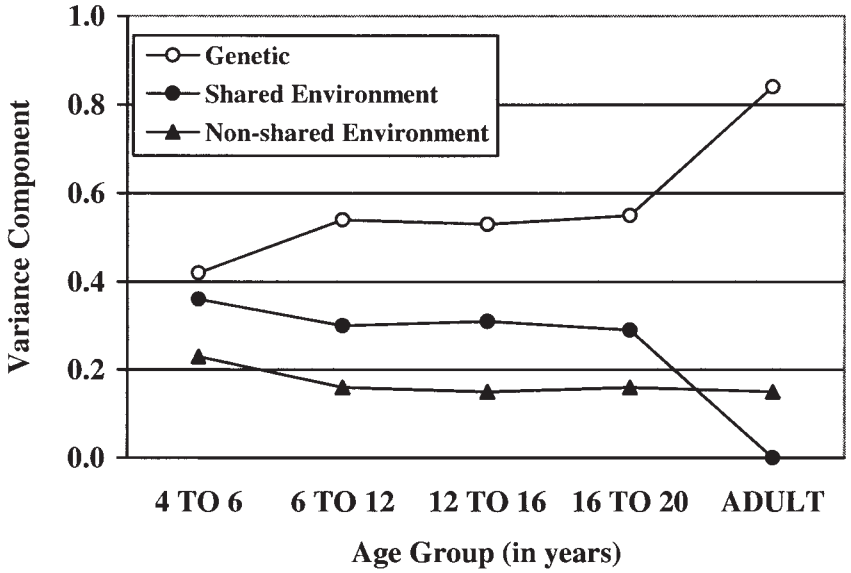

Figure 3. Estimated proportions of the total IQ variance attributable to genetic and environmental (shared and nonshared) effects. Note that only the nonshared (or within-family) environmental variance remains relatively constant across the entire age range. From "Behavioral Genetics of Cognitive Ability: A Life-Span Perspective," by M. McGue, T. J. Bouchard, Jr., W. G. Iacono, and D. T. Lykken, in R. Plomin and G. E. McClearn (Eds.), *Nature, Nurture, and Psychology* (p. 64), edited by R. Plomin and G. E. McClearn, 1993, Washington, DC: American Psychological Association. Copyright 1993 by the American Psychological Association. Reprinted with permission.

These results are corroborated by studies of monozygotic twins reared apart and of other kinships groups (Plomin et al., 2001). Because the variables usually proposed to explain mean racial-group differences are part of the shared family environment (such as social class, religious beliefs, cultural practices, father absence, and parenting styles), and these account for little variance *within* a race, they are unlikely to account for the differences *between* races. Rather, mean differences between races are primarily due to nonshared family effects, which include not only genetics but also a range of idiosyncratic environmental events that, within-families, affect one sibling and not the other (Jensen, 1997).

Hereditarians have also examined the question of whether group differences occur in shared and in nonshared environmental effects as well as in genetic effects. For example, Rushton and Osborne (1995) reanalyzed 125 Black and 111 White pairs of 12- to 18-year-old twins from the Georgia Twin Study and estimated their cranial capacities from head size measures. They found a lower range of heritabilities for Blacks than for Whites (12% to 31% against 47% to 56%) and a higher range of common environmental (i.e., shared family) effects for Blacks than for Whites (42% to 46% against 28% to 32%). However, these percentage differences between Blacks and Whites were not significant, although all heritabilities within each race were significantly above zero.

Also relevant to the question of the mean Black–White group differences are the changes in heritability that occur with increases in age (see Plomin et al., 2001). The average correlation of IQ between full siblings reared together reaches .49 in adulthood. The correlation in IQ for siblings reared apart as children is .24, which increases to .49 in adulthood. This shows that siblings grow more similar to each other as they age. In genetically unrelated people reared together, such as adopted children, the correlation for IQ is .25 in childhood but decreases to .01 in adulthood (McGue et al., 1993). This shows, conversely, that unrelated people reared together grow less similar over time. Between childhood and adulthood the influence of the shared home environment on IQ decreases, whereas the effect of genetic similarity increases.

The diminishing or even vanishing effect of differences due to the shared home environment can best be understood in terms of three components of gene–environment correlation and the change in their relative importance during development (Plomin et al., 1977; Plomin et al., 2001). The *passive* component of the gene–environment correlation reflects all those things that happen to the phenotype, independent of its own characteristics. For example, children of academically oriented parents may inherit genes for academic ability and also be exposed (through no effort of their own) to stimulating intellectual environments. The *reactive* component of gene–environment correlation results from the reaction of others to the expression of genetically based abilities, as when children with an unusual curiosity about science are given chemistry sets. The *active* component of the gene–environment correlation results from children actively seeking experiences compatible with their genotypes, for example, going to science fairs rather than sports events or music concerts. From early childhood to late adolescence the predominant component of the gene–environment covariance gradually shifts from passive to reactive to active. The child's enlarging world is like a cafeteria in which choices become increasingly biased by genetic factors (Scarr, 1996; Scarr & McCartney, 1983). As individuals mature they seek out and even create their own experiential environment.

Section 6: Race, Brain Size, and Cognitive Ability

Studies on over 700 participants show that individuals with larger brain volumes have higher IQ scores. About two dozen studies using magnetic resonance imaging (MRI) to measure the volume of the human brain have found an overall correlation with IQ of greater than .40 (Rushton & Ankney, 1996; P. A. Vernon, Wickett, Bazana, & Stelmack, 2000). The correlation of .40 using MRI is much higher than the .20 correlation found in earlier research using simple head size measures, although the .20 correlation is also reliable and significant. Rushton and Ankney (1996) reviewed 32 studies correlating measures of external head size with IQ scores or with measures of educational and occupational achievement, and they found a mean $r = .20$ for people of all ages, both sexes, and various ethnic backgrounds, including African Americans.

The most likely reason why larger brains are, on average, more intelligent than smaller brains is that they contain more neurons and synapses, which make them more efficient. Haier et al. (1995) tested the brain efficiency hypothesis by using MRI to measure brain volume and glucose metabolic rate to measure

glucose uptake (an indicator of energy use). They found a correlation of –.58 between glucose metabolic rate and IQ, suggesting that more intelligent individuals have more efficient brains because they use less energy in performing a given cognitive task. Several other studies supporting the brain-size/efficiency model were reviewed in Gignac, Vernon, and Wickett (2003). In any individual, however, energy use increases with the increasing complexity of the cognitive task.

Estimates from twin studies indicate that genes contribute from 50% to 90% of the variance to both cranial capacities based on external head size measures and to brain volume measured by MRI (Bartley, Jones, & Weinberger, 1997; Pennington et al., 2000; Posthuma et al., 2002; Rushton & Osborne, 1995; Thompson et al., 2001). Common genetic effects mediate from 50% to 100% of the brain-size/IQ correlation (Pennington et al., 2000; Posthuma et al., 2002). Studies have also shown that correlations between brain size and IQ also hold true *within* families as well as *between* families (Gignac et al., 2003; Jensen, 1994; Jensen & Johnson, 1994), which also implies shared genetic effects. However, one study that examined only sisters failed to find the within-family relation (Schoenemann, Budinger, Sarich, & Wang, 2000). Families with larger brains overall tend to have higher IQs and, within a family, the siblings with the larger brains tend to have higher IQ scores. The within-family finding is of special interest because it controls for most of the sources of variance that distinguish families, such as social class, styles of child rearing, and general nutrition, that differ between families.

Race differences in average brain size are observable at birth. A study by Rushton (1997) analyzed recorded head circumference measurements and IQ scores from 50,000 children in the Collaborative Perinatal Project followed from birth to age 7 (Broman, Nichols, Shaugnessy, & Kennedy, 1987). Using the head circumference measures to calculate cranial capacity at birth, 4 months, 1 year, and 7 years, at each of these ages, the Asian American children averaged larger cranial volumes than did the White children, who averaged larger cranial volumes than did the Black children. Within each race, cranial capacity correlated with IQ scores. By age 7, the Asian American children averaged an IQ of 110; the White children, 102; and the Black children 90. Because the Asian American children were the shortest in stature and the lightest in weight while the Black children were the tallest in stature and the heaviest in weight, these average race differences in brain-size/IQ relations were not due to body size.

External head size measurements (length, width, height) also have been used to estimate cranial capacities in adults. Rushton carried out five studies of large archival data sets. The first (Rushton, 1991) examined head size measures in 24 international military samples collated by the U.S. National Aeronautics and Space Administration. After adjusting for the effects of body height, weight, and surface area, the mean cranial capacity for East Asians was 1,460 cm^3 and for Europeans 1,446 cm^3. The second (Rushton, 1992) demonstrated that even after adjusting for the effects of body size, sex, and military rank in a stratified random sample of over 6,000 U.S. Army personnel, the average cranial capacity of East Asians, Whites, and Blacks were 1,416, 1,380, and 1,359 cm^3, respectively. The third study (Rushton, 1993) reanalyzed a set of anthropometric data originally published by Melville Herskovits (who concluded there were not race differences in cranial capacity) and found Whites averaged a cranial capacity of 1,421 and

Blacks, 1,295 cm^3. The fourth study (Rushton, 1994) analyzed data obtained on tens of thousands of people from around the world collated by the International Labor Office in Geneva, Switzerland. It found that after adjusting for the effects of body size and sex, samples from the Pacific Rim, Europe, and Africa had average cranial capacities, of 1,308, 1,297, and 1,241 cm^3 respectively. Finally, Rushton and Osborne (1995) analyzed the Georgia Twin Study of adolescents and found that after correcting for body size and sex, Whites had an average cranial capacity of 1,269 cm^3, Blacks 1,251 cm^3.

Rushton's results, based on calculating average cranial capacity from external head size measures, join those from dozens of other studies from the 1840s to the present on different samples using three different methods (endocranial volume from empty skulls, wet brain weight at autopsy, and high-tech MRI). All show the same strong pattern of East Asians averaging larger and heavier brains than Whites who average larger and heavier brains than Blacks. For example, using MRI technology, Harvey, Persaud, Ron, Baker, and Murray (1994) found that 41 Blacks in Britain averaged a smaller brain volume than did 67 British Whites.

The American anthropologist Samuel George Morton (1849) filled over 1,000 skulls with packing material to measure endocranial volume and found that Blacks averaged about 5 cubic inches less cranial capacity than Whites. His results were confirmed by Todd (1923), H. L. Gordon (1934), and Simmons (1942). The most extensive study of race differences in endocranial volume to date measured 20,000 skulls from around the world and reported East Asians, Europeans, and Africans had average cranial volumes of 1,415, 1,362, and 1,268 cm^3, respectively (Beals, Smith, & Dodd, 1984).

Using the method of weighing brains at autopsy, Paul Broca (1873) reported that Whites averaged heavier brains than did Blacks, with larger frontal lobes and more complex convolutions. (Broca also used endocranial volume and found East Asians averaged larger cranial capacities than Europeans, who averaged larger than Blacks.) Other early autopsy studies found a mean Black–White group difference in brain weight of about 100 g (Bean, 1906; Mall, 1909; Pearl, 1934; Vint, 1934). A more recent autopsy study of 1,261 American adults found that the brains of 811 White Americans in their sample averaged 1,323 g and the brains of 450 Black Americans averaged 1,223 g—a difference of 100 g (Ho, Roessmann, Straumfjord, & Monroe, 1980). Because the Blacks and Whites in the study were similar in body size, this was not responsible for the differences in brain weight.

Rushton (2000; Rushton & Ankney, 1996) summarized the world database using the three methods on which there are a sufficient number of studies (autopsies, endocranial volume, and head measurements), as well as head measurements corrected for body size (see Rushton, 2000, pp. 126–132, Table 6.6). The results in cm^3 or equivalents were as follows: East Asians = 1,351, 1,415, 1,335, and 1,356 (M = 1,364); Whites = 1,356, 1,362, 1,341, and 1,329 (M = 1,347); and Blacks = 1,223, 1,268, 1,284, and 1,294 (M = 1,267). The overall mean for East Asians is 17 cm^3 more than that for Whites and 97 cm^3 more than that for Blacks. Within-race differences due to differences in method of estimation averaged 31 cm^3. Because 1 cubic inch of brain matter contains millions of brain cells and hundreds of millions of synapses or neural connections, these group differences in average brain size may explain group differences in average IQ.

Jensen and Johnson (1994) showed that for both Blacks and Whites, the head size by IQ correlation is true within families as well as between families, indicating the intrinsic or functional relationship mentioned earlier. Further, within each sex, Blacks and Whites fit the same regression line of head size on IQ. When Blacks and Whites are perfectly matched for true-score IQ (i.e., IQ corrected for measurement error) at either the Black mean or the White mean, the overall average Black–White group difference in head circumference is virtually nil. (Matching Blacks and Whites for IQ eliminates the average difference in head size, but matching the groups on head size does not equalize their IQs. This is what one would expect if brain size is only one of a number of brain factors involved in IQ.)

In another analysis of the Georgia Twin Study, Jensen (1994) showed that the mean Black–White group difference in head/brain size is also related to the magnitude of the mean Black–White group difference in *g*. The correlation coefficient of each test with the head measurements was correlated with the magnitude of the Black–White group difference on that test, thus forming two vectors. The column vector of IQ test and head size correlations indicated a correlation of .51 ($p < .05$) with the vector of standardized Black–White group differences on each of the tests.

Section 7: Mean Race–IQ Differences and Transracial Adoption Studies

"Transracial adoption is the human analog of the cross-fostering design, commonly used in animal behavior genetics research. . . .There is no question that adoption constitutes a massive intervention" (Scarr & Weinberg, 1976, p. 726). Studies of Korean and Vietnamese children adopted into White homes show that although as babies many had been hospitalized for malnutrition, they nonetheless grew to have IQs 10 or more points higher than their adoptive national norms. By contrast, Black and mixed-race (Black–White) children adopted into White middle-class families typically have lower average scores than the White siblings with whom they had been reared or than White children adopted into similar homes.

The Minnesota Transracial Adoption Study, the largest and best-known transracial study, was designed specifically by Sandra Scarr and Richard Weinberg to separate genetic factors from rearing conditions as causal influences on the cognitive performance of Black children (Scarr & Weinberg, 1976; Weinberg, Scarr, & Waldman, 1992). It is also the only transracial adoption study that includes a longitudinal follow-up, with testing at ages 7 and 17 years. Scarr and Weinberg compared the IQ and academic achievement scores of Black, White, and mixed-race Black/White children adopted into upper-middle-class White families in Minnesota by adopting parents whose mean IQ was more than 1 standard deviation above the population mean of 100 (see Table 2). The biological children of these parents were also tested.

The first testing of 265 children was carried out in 1975 when they were 7 years old and the second in 1986 when the 196 remaining in the study were 17 years old. The 7-year-old White biological (i.e., nonadopted) children had an average IQ of 117 (see Table 2, 2nd column), similar to that found for children of White upper-middle-class parents. The adopted children with two White biological parents had a mean IQ of 112. The adopted children with one Black and one

Table 2
Comparison of Cognitive Performance Measures at Ages 7 and 17 in Biological and Adopted (White, Mixed-Race, and Black) Children, All Reared in Middle-Class White Families

Children's background	Age 7 IQ	Age 17 IQ	Age 17 grade point average	Age 17 class rank (percentile)	Age 17 school aptitude (percentile)[a]
Biological parents	120	115	—	—	—
Nonadopted, with two White biological parents (*N* at age 7 = 143; *N* at 17 = 104)	117	109	3.0	64	69
Adopted, with two White biological parents (*N* at age 7 = 25; *N* at 17 = 16)	112	106	2.8	54	59
Adopted, with one White and one Black biological parent (*N* at age 7 = 68; *N* at 17 = 55)	109	99	2.2	40	53
Adopted, with two Black biological parents (*N* at age 7 = 29; *N* at 17 = 21)	97	89	2.1	36	42

Note. From "The Minnesota Transracial Adoption Study: A Follow-Up of IQ Test Performance at Adolescence," by R. A. Weinberg, S. Scarr, and I. D. Waldman, 1992, *Intelligence, 16,* pp. 117–135. Copyright 1992 by Ablex Publishing. Adapted with permission.
[a] Based on national norms (weighted mean of four percentiles).

White biological parent averaged 109. The adopted children with two Black biological parents had an average IQ of 97. (A mixed group of 21 Asian, North American Indian, and Latin American Indian adopted children averaged an IQ of 100 but were not included in the main statistical analyses.)

Scarr and Weinberg (1976) interpreted the results of the testing at age 7 as support for the culture-only position. They drew special attention to the fact that the mean IQ of 105 for all "socially classified" Black children (i.e., those with either one or two Black parents) was significantly above the U.S. White mean. The poorer performance of children with two Black biological parents was attributed to their more difficult and later placement. Scarr and Weinberg also pointed out that this latter group had both natural and adoptive parents with somewhat lower educational levels and abilities (2 points lower in adoptive parents' IQ). They found no evidence for the *expectancy effects* hypothesis that adoptive parents' beliefs about the child's racial background influence the child's intellectual development. The mean score for 12 children wrongly believed by their adoptive parents to have two Black biological parents was virtually the same as that of the 56 children correctly classified by their adoptive parents as having one Black and one White biological parent.

Table 2 also presents the results for the 196 children retested at age 17 (Weinberg et al., 1992). There were four independent assessments of the children's cognitive performance at this later age: (a) an individually administered IQ test, (b) an overall grade point average, (c) a class rank based on school performance, and (d) four special aptitude tests in school subjects administered by the educational authority, which we averaged. The results are concordant with the earlier testing. The nonadopted White children had a mean IQ of 109, a grade point average of 3.0, a class rank at the 64th percentile, and an aptitude score at the 69th percentile. The adopted children with two White biological parents had a mean IQ of 106, a grade point average of 2.8, a class rank at the 54th percentile, and an aptitude score at the 59th percentile. The adopted children with one Black and one White biological parent had a mean IQ of 99, a grade point average of 2.2, a class rank at the 40th percentile, and an aptitude score at the 53rd percentile. The adopted children with two Black biological parents had a mean IQ of 89, a grade point average of 2.1, a class rank at the 36th percentile, and an aptitude score at the 42nd percentile. (The 12 remaining mixed group of Amerindian/Asian children had an IQ of 96 with no data provided on school achievement.)

Because different tests based on different standardization groups were used in the first testing than in the follow-up, the overall average difference of about 8 IQ points (evident for all groups, including the nonadopted group) between the two test periods does not bear on the hypothesis of interest. The relevant comparisons are those between the adopted groups of different races *within* each age level. The mean of 89 for adopted children with two Black parents was slightly above the national Black mean of 85 but not above the Black mean for Minnesota.

Weinberg et al. (1992) interpreted their follow-up results as further support for the culture-only theory. Emphasizing the beneficial effects of the rearing environment, they pointed out that at both age 7 and 17 all groups of adopted children averaged above their expected population means. Their analyses frequently combined the two "socially classified Black" groups with "other" mixed-race children who had one parent of unknown, Asian, Amerindian, or other racial

background. In their age 17 breakdowns, Weinberg et al. (1992, p. 132) stated that "[b]iological mothers' race remained the best single predictor of adopted child's IQ when other variables were controlled," which they then attributed to "unmeasured social characteristics." Their overall conclusion was that "the social environment maintains a dominant role in determining the average IQ level of Black and interracial children and that both social and genetic variables contribute to individual variations among them" (p. 133).

Levin (1994) and Lynn (1994) disputed Weinberg et al.'s (1992) culture-only interpretation. They each proposed a straightforward, hereditarian alternative: The mean IQ and school achievement scores of Black children reflected their degree of African ancestry. At both age 7 and 17, the adopted children with two Black biological parents had lower average IQs and school achievement scores than did those with one Black and one White biological parent, and these children, in turn, averaged lower scores than did those with two White biological parents. Waldman, Weinberg, and Scarr (1994) responded to Levin (1994) and Lynn (1994) with further regression analyses that indicated the children's preadoptive experience was confounded with racial ancestry, and so an unambiguous interpretation of the results was not possible.

Subsequently, Jensen (1998b) discussed these studies at length and reviewed the evidence showing that age of adoption does not influence children's IQ scores after age 7 (e.g., Fisch, Bilek, Deinard, & Chang, 1976). Studies of severely malnourished, late-adopted, East Asian children (see below) provide substantial evidence that age of adoption does not adversely influence IQ in transracial adoptions. More generally, as reviewed in Section 5, dozens of adoption, twin, and family studies of Whites show that although the shared-family environmental component of true-score IQ variance can be quite large at age 7, by late adolescence it is the smallest component. After that age, genetic and within-family (nonshared) environmental effects account for the largest components of variance (see Figure 3).

Small sample studies of very young children reared in enriched environments sometimes find an absence of the usual race differences in IQ. In two studies of 2- to 5-year-olds raised in English residential nurseries, Tizard (1974) compared Black (African and West Indian), White, and mixed-parentage children and found no significant differences among the three groups on several language comprehension tests and on the Wechsler Preschool and Primary Scale of Intelligence (WPPSI); the single significant difference was in favor of the non-White children. Moore (1986) found that at age 7, 23 Black children adopted by middle-class *White* families had a mean IQ of 117, whereas a similar group of children adopted by middle-class *Black* families had a mean IQ of 104, both significantly above the national Black mean of 85. To be more informative, future studies need to be supplemented by follow-up testing, as in the Minnesota Study. Behavior genetic studies consistently show that, as people age, their genes exert ever more influence, whereas family socialization effects decrease (see Figure 3). Trait differences not apparent early in life begin to appear at puberty and are completely apparent by age 17.

Three studies of East Asian children adopted by White families support the hereditarian hypothesis. In the first, 25 four-year-olds from Vietnam, Korea, Cambodia, and Thailand, all adopted into White American homes prior to 3 years

of age, excelled in academic ability with a mean IQ score of 120, compared with the U.S. norm of 100 (Clark & Hanisee, 1982). Prior to placement, half of the babies had required hospitalization for malnutrition.

In the second study, Winick, Meyer, and Harris (1975) found 141 Korean children adopted as infants by American families exceeded the national average in both IQ and achievement scores when they reached 10 years of age. The principal interest of the investigators was on the possible effects of severe malnutrition on later intelligence, and many of these Korean children had been malnourished in infancy. When tested, those who had been severely malnourished as infants obtained a mean IQ of 102; a moderately well-nourished group obtained a mean IQ of 106; and an adequately nourished group obtained a mean IQ of 112.

A study by Frydman and Lynn (1989) examined 19 Korean infants adopted by families in Belgium. At about 10 years of age, their mean IQ was 119, the verbal IQ was 111, and the performance IQ was 124. Even correcting the Belgian norms upward to 109 to account for the increase in IQ scores over time (about 3 IQ points a decade; see Section 13), the Korean children still had a statistically significant 10-point advantage in mean IQ over indigenous Belgian children. Neither the social class of the adopting parents nor the number of years the child spent in the adopted family had any effect on the child's IQ.

Section 8: Mean Race–IQ Differences and Racial Admixture

In the Minnesota Transracial Adoption Study, the IQs of the mixed-race (Black/White) adoptees averaged between those of the "nonmixed" White and the "nonmixed" Black adoptees, as expected under a genetic hypothesis (see Table 2). Results from some other types of studies are also consistent with that hypothesis. In her review, Shuey (1966) found that in 16 of 18 studies in which skin color could be used as a proxy for amount of admixture, Blacks with lighter skin color averaged higher scores than those with darker skin, although the magnitude of the association was quite low ($r = .10$). The Black American average IQ of 85 (15 points higher than the sub-Saharan African average of 70; see Section 3) is also consistent with the genetic hypothesis, given the approximately 20% White admixture of Black Americans (Chakraborty, Kamboh, Nwankwo, & Ferrell, 1992; Parra et al., 1998). The mixed-race "Colored" population of South Africa also has an average IQ of 85, intermediate to the respective African and White means of 70 and 100 (Owen, 1992). Early studies of brain weight data also fit with the genetic hypothesis. Bean (1906) found, as did Pearl (1934), that the greater the amount of White admixture (judged independently from skin color), the higher the mean brain weight at autopsy in Black groups. More recent data of this nature are not available.

The average IQ scores of around 70 for Black Americans in certain areas of the Deep South of the United States where the degree of White admixture is significantly below the general average (Chakraborty et al., 1992; Parra et al., 1998) are also consistent with the hereditarian interpretation of the effects of hybridization. An average IQ of 71 was found for all of the Black children in an entire school district from a rural county in Georgia; the average White IQ in the same county was 101 (Jensen, 1977). Similarly, Stanley and Porter (1967) found the scores on the SAT of all-Black college students in Georgia were too low to be

predictive of college grades, thereby raising the question of whether test scores on Black Americans are as valid as those for White Americans. However, when Hills and Stanley (1970) gave the School and College Ability Test (a much easier test to pass) to similar students, they found that their scores were normally distributed and *did* predict college grades, though the average for the Black college students was at about the 50th percentile on eighth-grade national norms.

Most recently, Lynn (2002) and Rowe (2002) analyzed data from large, publicly available, archival data sets, which show that groups of mixed-race individuals have mean scores intermediate to unmixed groups of Blacks and of Whites. Lynn examined the 1982 National Opinion Research Center's survey of a representative sample of the adult population, excluding non-English speakers. The 442 Blacks in the sample were asked whether they would describe themselves as "very dark," "dark brown," "medium brown," "light brown," or "very light." The correlation between these self-ratings and a 10-word vocabulary test score was .17 ($p < .01$). Rowe examined the 1994 National Longitudinal Study of Adolescent Health's survey of a representative sample of youths, with intentional oversampling of Black children of highly educated parents. The mean age for the entire sample (9,830 Whites, 4,017 Blacks, and 119 mixed-race individuals) was 16 years. The Black adolescents averaged a lower birth weight, a lower verbal IQ, and a higher number of sexual partners than did the White adolescents. For each characteristic, the mixed-race mean fell between the means of the other two groups. Rowe found the social class explanation of the group differences "unconvincing" because, of the three variables, only verbal IQ showed a moderate correlation with social class and statistically adjusting for it left the main findings unchanged. He also rejected the "discrimination based on skin tone" hypothesis because it was eliminated by deliberately selecting only those mixed-race adolescents who were judged by their interviewers to be Black, based on their physical appearance.

Three studies of racially mixed individuals at first appear to support the culture-only hypothesis against the genetic hypothesis. Eyferth (1961; Eyferth, Brandt, & Hawel, 1960) reported IQ data for out-of-wedlock children fathered by soldiers stationed in Germany after World War II and then reared by White German mothers. The mean IQs for 83 White children and for 98 racially mixed children were both about 97 (97.2 for the Whites, 96.5 for the racially mixed). As Loehlin et al. (1975, pp. 126–128) noted, however, these results are ambiguous for three reasons. First, the children were still very young when tested. One third of the children were between 5 and 10 years of age, and two thirds were between 10 and 13 years. As discussed in Section 5 (see Figure 3), behavior genetic studies show that while family socialization effects on IQ are often strong before puberty, after puberty they dwindle, sometimes to zero. Second, 20% to 25% of the "Black" fathers were not African Americans but French North Africans (i.e., largely Caucasian or "Whites" as we have defined the terms here). Third, there was rigorous selection based on IQ score in the U.S. Army at the time, with a rejection rate for Blacks on the preinduction Army General Classification Test of about 30%, compared with 3% for Whites (see Davenport, 1946, Tables I and III).

The second study reports a 9-point IQ advantage for the 4-year-old offspring of couples with a White mother and a Black father (mean IQ = 102, $N = 101$) compared with those from the offspring of a Black mother and a White father

(mean IQ = 93, N = 28). Willerman, Naylor, and Myrianthopoulos (1974), assuming White mothers provide better pre- or postnatal environments for their children than do Black mothers, interpreted their data as more consistent with a cultural than a genetic hypothesis (see also Nisbett, 1998). However, Loehlin et al. (1975, p. 126) noted that the mixed-race pairs with White mothers averaged almost a year more schooling than did the pairs with Black mothers. Thus the White mothers may have had a higher average IQ than the Black ones. The mid-parent IQs have to be the same for the results to be interpretable. Also, the two sets of mixed-race children averaged an IQ of 98, intermediate to the White and Black children in the sample from whom the mixed-race children had been drawn (IQs = 105 and 91, respectively; Broman, Nichols, & Kennedy, 1975, p. 43).

The third study seeming to support the culture-only hypothesis is a subsidiary analysis by Moore (1986; see Section 8) of a small number of 7-year-old children adopted by middle-class White parents. Moore found no difference in IQ between those children with only one and those with two Black biological parents. The mean IQ for the group of 9 adopted children with two Black biological parents was 109, and the mean IQ for the group of 14 children with one Black and one White biological parent was 107. Given the young age of these children, a follow-up to adolescence would be informative.

Studies of blood groups provide no support for the hereditarian perspective. Both Loehlin, Vandenberg, and Osborne (1973) and Scarr, Pakstis, Katz, and Barker (1977) found that blood groups distinguishing African from European ancestry did not predict IQ scores in Black samples. However, these studies failed to choose genetic markers with large allele frequency differences between Africans and Europeans (Jensen, 1998b, pp. 480, 524 n.64).

Molecular genetic technology was unsophisticated in the 1970s. In the future, the issue may be resolved by calculating individual admixture through the use of DNA markers as already occurs in medicine (Risch et al., 2002). On the basis of existing surveys, an individual's racial group can be determined by testing his or her DNA at 100 random sites along the genome, or at 30 specifically chosen ones. Even different ethnic groups within a race can be distinguished using some 50 specifically chosen sites. A genetic hypothesis predicts that for those Black individuals who possess more White genes, their physical, behavioral, and other characteristics will approach those of Whites.

Although the studies of racial hybrids are generally consistent with the genetic hypothesis, to date they are not conclusive. It may be true, for example, that lighter skinned Cape Coloreds and African Americans have better nutrition, have greater opportunities for learning, or are treated better by their societies. On the other hand, the Minnesota Transracial Adoption Study (Table 2) held many such factors constant and removed the most frequently proposed causal agents such as poverty, malnutrition, poor schools, and dysfunctional neighborhoods. Yet, here too, the mixed-race children had a higher mean IQ than did the children of two Black parents, and the means for each group were very similar to those for their respective counterparts elsewhere in the United States. The discussion in this section is particularly supportive of Loehlin's (2000) conclusion that "Research using larger samples and better techniques for estimating ancestry is called for and quite feasible" (p. 188).

Section 9: Mean Race–IQ Differences and Regression to the Mean

Regression toward the mean provides still another method of testing if the group differences are genetic. Regression toward the mean is seen, on average, when individuals with high IQ scores mate and their children show lower scores than their parents. This is because the parents pass on some, but not all, of their genes to their offspring. The converse happens for low IQ parents; they have children with somewhat higher IQs. Although parents pass on a random half of their genes to their offspring, they cannot pass on the particular combinations of genes that cause their own exceptionality. This is analogous to rolling a pair of dice and having them come up two 6′s or two 1′s. The odds are that on the next roll, you will get some value that is not quite as high (or as low). Physical and psychological traits involving dominant and recessive genes show some regression effect. Genetic theory predicts the magnitude of the regression effect to be smaller the closer the degree of kinship between the individuals being compared (e.g., identical twin > full-sibling or parent–child > half-sibling). Culture-only theory makes no systematic or quantitative predictions.

For any trait, scores should move toward the average for that population. So in the United States, genetic theory predicts that the children of Black parents of IQ 115 will regress toward the Black IQ average of 85, whereas children of White parents of IQ 115 will regress toward the White IQ average of 100. Similarly, children of Black parents of IQ 70 should move up toward the Black IQ average of 85, whereas children of White parents of IQ 70 should move up toward the White IQ average of 100. This hypothesis has been tested and the predictions confirmed. Regression would explain why Black children born to high IQ, wealthy Black parents have test scores 2 to 4 points lower than do White children born to low IQ, poor White parents (Jensen, 1998b, p. 358). High IQ Black parents do not pass on the full measure of their genetic advantage to their children, even though they gave them a good upbringing and good schools, often better than their own. (The same, of course, applies to high IQ White parents.) Culture-only theory cannot predict these results but must argue that cultural factors somehow imitate the effect theoretically predicted by genetic theory, which have also been demonstrated in studies of physical traits and in animals.

Jensen (1973, pp. 107–119) tested the regression predictions with data from siblings (900 White sibling pairs and 500 Black sibling pairs). These provide an even better test than parent–offspring comparisons because siblings share very similar environments. Black and White children matched for IQ had siblings who had regressed approximately halfway to their respective population means rather than to the mean of the combined population. For example, when Black children and White children were matched with IQs of 120, the siblings of Black children averaged close to 100, whereas the siblings of White children averaged close to 110. A reverse effect was found with children matched at the lower end of the IQ scale. When Black children and White children are matched for IQs of 70, the siblings of the Black children averaged about 78, whereas the siblings of the White children averaged about 85. The regression line showed no significant departure from linearity throughout the range of IQ from 50 to 150, as predicted by genetic theory but not by culture-only theory.

Section 10: The Race–Behavior Matrix

Around the world, the rate of dizygotic (i.e., two-egg) twinning is less than 4 per 1,000 births among East Asians, 8 among Whites, and 16 or greater among Blacks (Bulmer, 1970). Multiple birthing rates have been shown to be heritable, based on the race of the mother, regardless of the race of the father, as found in East Asian–White crosses in Hawaii and White–Black crosses in Brazil (Bulmer, 1970).

On average, Black babies are born a week earlier than White babies, yet they are more mature as measured by pulmonary function, amniotic fluid, and bone development. In the United States, 51% of Black children have been born by week 39 of pregnancy compared with 33% of White children. Black African babies, even those born to mothers in the professional classes, are also born earlier than White babies (Papiernik, Cohen, Richard, de Oca, & Feingold, 1986). They are not born premature but sooner, and they are biologically more mature.

After birth, Black babies continue to mature faster, on average, than White babies, whereas East Asian babies average an even slower rate. X-rays show a faster rate of average bone growth in Black children than in White children, and a faster rate in White children than in East Asian children (Eveleth & Tanner, 1990, pp. 154–155). Black babies at a given age also average greater muscular strength and a more accurate reach for objects. Black children average a younger age of sitting, crawling, walking, and putting on their own clothes than Whites or East Asians. The average age of walking is 13 months in East Asian children, 12 months in White children, and 11 months in Black children (Bayley, 1965; Brazelton & Freedman, 1971).

Blacks average a faster rate of dental development than do Whites, who have a faster rate than do East Asians. On average, Black children begin the first stage of permanent tooth growth at about 5.8 years, whereas Whites and East Asians do not begin until 6.1 years (Eveleth & Tanner, 1990, pp. 158–161). Blacks also have an earlier age of sexual maturity than do Whites, who in turn have an earlier average age than do East Asians, whether measured by age of first menstruation, first sexual experience, or first pregnancy (Rushton, 2000, pp. 147–150).

Myopia (nearsightedness) is positively correlated with IQ and may be caused by extra myelinization in the eye and so possibly linked to brain size (Miller, 1994). The relationship appears to be pleiotropic (Cohn, Cohn, & Jensen, 1988); that is, a gene affecting one trait also has some effect on one or more others. There are significant racial and ethnic differences in the frequency of myopia, with the highest rates found in East Asians, the lowest rates among Blacks, with Whites intermediate (Post, 1982).

Not just in the United States but around the world, East Asians and Blacks fall at the two ends of a continuum with Whites intermediate, not only on mean cognitive test scores and brain size measures but also on 60 life-history variables that provide measures of maturation, personality, reproduction, and social organization. It seems unlikely that social factors alone could produce this consistent pattern on so diverse a set of behaviors (see Table 3; Rushton, 2000, p. 5, Table 1.1 for complete list). This evidence raises the theoretical question of whether single traits such as intelligence are part of a broader "life-history" perspective.

Table 3
Worldwide Average Differences Among Blacks, Whites, and East Asians

Trait	Blacks	Whites	East Asians
Intelligence			
IQ test scores	85	102	106
Decision times	Slower	Intermediate	Faster
Cultural achievements	Low	High	High
Brain size			
Cranial capacity (cm^3)	1,267	1,347	1,364
Cortical neurons (millions)	13,185	13,665	13,767
Maturation rate			
Gestation time	Shorter	Longer	Longer
Skeletal development	Earlier	Intermediate	Later
Motor development	Earlier	Intermediate	Later
Dental development	Earlier	Intermediate	Later
Age of first intercourse	Earlier	Intermediate	Later
Age of first pregnancy	Earlier	Intermediate	Later
Life span	Shortest	Intermediate	Longest
Personality			
Aggressiveness	Higher	Intermediate	Lower
Cautiousness	Lower	Intermediate	Higher
Impulsivity	Higher	Intermediate	Lower
Self-concept	Higher	Intermediate	Lower
Sociability	Higher	Intermediate	Lower
Reproduction			
Two-egg twinning (per 1,000 births)	16	8	4
Hormone levels	Higher	Intermediate	Lower
Sex characteristics	Larger	Intermediate	Smaller
Intercourse frequencies	Higher	Intermediate	Lower
Permissive attitudes	Higher	Intermediate	Lower
Sexually transmitted diseases	Higher	Intermediate	Lower
Social organization			
Marital stability	Lower	Intermediate	Higher
Law abidingness	Lower	Intermediate	Higher
Mental health	Lower	Intermediate	Higher

Note. From *Race, Evolution, and Behavior: A Life History Perspective* (p. 5), by J. P. Rushton, 2000, Port Huron, MI: Charles Darwin Research Institute. Copyright 2000 by J. P. Rushton. Adapted with permission.

Section 11: Mean Race–IQ Differences and Human Origins

The currently most commonly accepted view of human origins, the "Out-of-Africa" theory, posits that *Homo sapiens* arose in Africa about 150,000 years ago, expanded northward beyond Africa about 100,000 years ago, with a European–East Asian split about 41,000 years ago (Cavalli-Sforza et al., 1994; Stringer & McKie, 1996). In Cavalli-Sforza's (2000) maximum likelihood tree devised on the basis of molecular genetic markers, the most distant group was the Africans, with Europeans and Asians being closer. Cavalli-Sforza observed, "All world trees place the earliest split between Africans and non-Africans, which is expected given that all humans originated in Africa" (p. 72). This is also the conclusion of other reviewers (e.g., Risch et al., 2002).

Evolutionary selection pressures were different in the hot savanna where

Africans lived than in the cold northern regions Europeans experienced, or the even colder Arctic regions of East Asians. These ecological differences affected not only morphology but also behavior. It has been proposed that the farther north the populations migrated out of Africa, the more they encountered the cognitively demanding problems of gathering and storing food, gaining shelter, making clothes, and raising children successfully during prolonged winters (Rushton, 2000). As these populations evolved into present-day Europeans and East Asians, the ecological pressures selected for larger brains, slower rates of maturation, and lower levels of testosterone—with concomitant reductions in sexual potency, aggressiveness, and impulsivity; increases in family stability, advanced planning, self-control, rule following, and longevity; and the other characteristics listed in Table 3. The fact that the three-way pattern in IQ, brain size, and other traits is not unique to the United States but occurs internationally is consistent with a single, general (genetic–evolutionary) theory, whereas culture-only theory must invoke a number of highly localized, specific explanations.

As *Homo sapiens* migrated further away from Africa, the random genetic mutations that occur at a constant rate in all living species accumulated, along with the adaptive changes. The resulting differences in allele frequencies are sufficient for numerous and extensive genetic investigations to yield essentially the same picture and identify the same major racial groupings as did the morphological markers of classical anthropology. The greatest genetic divergence within the human species is between Africans (who have had the most time for random mutations to accumulate) and non-Africans (Cavalli-Sforza 2000; Cavalli-Sforza et al., 1994; Nei & Roychoudhury, 1993). Jensen (1998b, pp. 517–520) carried out a principal-components analysis of data on genetic markers from Nei and Roychoudhury (1993) and found the familiar clustering of races: (a) East Asians, (b) Europeans and East Indians, (c) South Asians and Pacific Islanders, (d) Africans, (e) North and South Amerindians and Eskimos, and (f) Aboriginal Australians and Papuan New Guineans. Howells's (1993) analysis of between-groups variation in craniometric data also revealed a similar population tree. The genetic hypothesis is consistent with the latest findings on human origins and genetic variation, whereas culture-only theory is indifferent to them (Crow, 2002).

Section 12: How Well Have Culture-Only Theories of Mean Race–IQ Differences Held Up?

Culture-only hypotheses have not explained the mean Black–White group differences in IQ. (They have especially not explained the findings on East Asians.) One early view was that the mean Black–White group difference in IQ was due to the then obvious differences in (segregated) school facilities (Myrdal, 1944). However, despite the U.S. Supreme Court *Brown v. Board of Education* (1954) decision striking down segregated schooling, and the consequent nationwide program of school busing, the mean Black–White group difference has not decreased. Moreover, the Coleman Report (Coleman et al., 1966) found that the racial composition of schools per se was not related to achievement in either Blacks or Whites. Most of the variation in IQ scores occurred *within* schools and less than 20% occurred *between* schools. Negligible, and in some cases, *negative* correlations were found between IQ and variables such as pupil expenditure,

teachers' salaries, teachers' qualifications, student/teacher ratios, and the availability of other school professionals (see also Coleman, 1990–1991).

The most frequently stated culture-only hypothesis is that the mean IQ differences are due to SES. In fact, controlling for SES only reduces the mean Black–White group difference in IQ by about a third, around 5 IQ points. The genetic perspective does not regard this control for SES as being entirely environmental. It holds that the parents' socioeconomic level in part reflects their genetic differences in intelligence. Moreover, according to the culture-only theory, as Black groups advance up the socioeconomic ladder, their children should be less exposed to environmental deficits and therefore should do better and, by extension, close the distance separating the Black mean with the White. In fact, the magnitude of the mean Black–White group difference in IQ for higher SES levels, when measured in standard deviations, is larger (Herrnstein & Murray, 1994, pp. 286–289).

Other nongenetic hypotheses are that standard IQ tests are culturally biased because the test items are not equally familiar and motivating to all groups or that they only measure familiarity with middle-class language or culture. However, despite attempts to equate items for familiarity and culture-fairness, no "culture-fair" test has eliminated the mean group difference. American Blacks actually have higher average scores on culturally loaded tests than on culturally reduced tests, which is the opposite to what is found for some other groups such as Mexican Indians and East Asians. (The mean Black–White group differences are greatest on the *g* factor, regardless of the type of test from which *g* is extracted; see Section 4.) Moreover, the three-way pattern of mean Black–White–East Asian group differences occurs worldwide on culture-fair reaction time measures, which all children can do in less than 1 s (see Section 3).

Subsequent culture-only hypotheses have pointed to specific aspects of deprivation as possible determinants of IQ. These include the following: (a) lack of reading material in the home, (b) poor cultural amenities in the home, (c) weak structural integrity of the home, (d) foreign language in the home, (e) low preschool attendance, (f) no encyclopedia in the home, (g) low level of parental education, (h) little time spent on homework, (i) low parental educational desires for child, (j) low parental interest in school work, (k) negative child self-concept (self-esteem), and (l) low child interest in school and reading. However, both within-race kinship studies and across-race adoption studies show that these environmental variables have increasingly smaller effects on the adoptees' IQ as they reach adolescence (see Sections 5 and 7). Moreover, other studies found that American Indians and East Asians averaged higher in IQ than Blacks, even though they averaged lower on these proposed causal factors (Coleman et al., 1966, p. 20). Another example comes from the Inuit, who live above the Arctic Circle and have higher average IQs than do either American or Jamaican Blacks (Berry, 1966; MacArthur, 1968) even though their socioeconomic conditions are extremely poor and unemployment is high (P. E. Vernon, 1965, 1979).

In the 1960s, culture-only theory formed the basis for implementing "Head Start"-type intervention programs as a way to eliminate the group differences in IQ and scholastic achievement. Although federal matching grants were given to improve the learning skills, social skills, and health status of low-income preschool children so that they could begin schooling on an equal footing with their

more advantaged peers, the mean Black–White group difference in IQ was not eliminated or permanently reduced. Currie and Thomas (1995) reviewed the literature and carried out a longitudinal study using a national sample of over 4,000 children in which they compared siblings to control for selection bias. They found that although Head Start led to large and significant *immediate* gains in test scores for both White and Black groups, these gains were quickly lost for Black groups, although some remained for White groups. Even more intensive and prolonged educational interventions than Head Start have not produced lasting effects on IQ or scholastic performance (Jensen, 1998b, pp. 333–344) or that generalize to other measures or criteria.

Some culture-only theorists propose that SES should not be assessed in terms of crude material measures but must be seen as a complex of attitudes, aspirations, self-images, and societal stereotypes (Loury, 2002; Ogbu, 2002; Sowell, 1994). Some of these types of cultural factors have been tested as well. Matching Black and White children for the geographical areas of their homes, the schools they attend, and other finer grade socioeconomic indicators again reduces the mean group IQ difference but does not eliminate it. Black children from the *best* areas and schools (those producing the highest average scores) still average slightly lower than do White children with the *lowest* socioeconomic indicators (Herrnstein & Murray, 1994, pp. 286–289; Jensen, 1998b, pp. 357–360). This is an anomaly for the culture-only theory but is explained by genetic theory through regression to the mean (see Section 10).

Other culture-only hypotheses have invoked Black role models, test anxiety, self-esteem, and racial stress as causal agents, but none of these have ever been consistently confirmed (Jensen, 1980, 1998b). Other ideas, such as stereotype threat (Steele, 1997), involuntary-minorities-are-castes (Ogbu, 2002), and race stigma (Loury, 2002), do not explain the low IQ of Africans south of the Sahara, where Blacks are in the majority. Nor is there any evidence from analyses of large archival data sets that unique minority-specific factors such as the history of slavery, White racism, lowered expectations, or heightened stress make cultural influences stronger for one group than for another (see Section 5). Neither can racial stigmatization (Loury, 2002) explain why East Asians average higher in IQ and brain size than Whites. A progressive theory of racial group differences must address *all* the known facts.

Culture-only theory must offer some explanation why its main variables—poverty, social class, religious beliefs, cultural practices, father absence, and parenting styles—account for so little variance *within* groups. Given these repeated findings, it is unlikely such variables can account for differences *between* groups (see Section 5). Adoption and twin studies show that the environmental variables influencing IQ and social behavior are primarily those that occur *within families* rather than *between families* (see Figure 3). Although the causes of within-group differences are logically separate from the causes of between-groups differences (Section 2), even when the combined set of within- and between-families variables is examined together, there are still no identifiable race-specific variables (Section 5).

It is always possible that new data with sharper hypotheses and better controls could require a revision of the finding of no shared family or minority-specific cultural effects on race–IQ differences. There were hints (but no more than that)

of a lower heritability and a greater shared environment component in Black adolescents than in White adolescents in Rushton and Osborne's (1995) twin study of cranial capacity (Section 5). Similarly, an epidemiological study of low-birth-weight and normal children, followed from 6 to 11 years of age, reported an IQ decline in mainly Black inner-city children with no similar IQ decline in mainly White suburban children. The authors interpreted their results as a between-community effect and the racial makeup of the schools the children attended, more than to individual and family factors (Breslau et al., 2001). Behavioral genetic designs using traditional modeling procedures (Section 5), along with new individual admixture measures on mixed-race participants (Section 8), could provide counterevidence to our conclusions. Unfortunately, behavioral geneticists (who have the most knowledge of the best techniques) have for the most part avoided the racial question.

One culture-only hypothesis currently enjoying much support is based on the secular increase in test scores, known as the *Flynn effect* because of the repeated demonstration by James Flynn (1984, 1987, 1999) that the average IQ in several countries has increased by about 3 points a decade over the last 50 years. Some have suggested that the Flynn effect implies that the 1 standard deviation difference in the mean Black–White IQ difference in the United States will gradually disappear over time (Flynn, 1999). However, one statistical analysis shows that the Flynn effect is not on the *g* factor, the principal source of the mean Black–White group difference.

Table 4 (based on Rushton, 1999) shows the results of a principal-components analysis of the secular gains in IQ from the United States, Germany, Austria, and Scotland, along with Black–White IQ difference scores from the United States, inbreeding depression scores from cousin-marriages in Japan, and *g* loadings from the standardization samples of the WISC–R and WISC–III. The relevant findings are as follows: (a) The IQ gains on the WISC–R and WISC–III form a cluster, showing that the secular trend in overall test scores is a reliable phenomenon; but (b) this cluster is *independent* of a second cluster formed by Black–White differences, inbreeding depression scores (a purely genetic effect), and *g* factor loadings (a largely genetic effect).

This analysis shows that the secular increase in IQ behaves differently from the mean Black–White group difference in IQ. Flynn's (1999) hypothesis that the IQ gains over time imply a purely environmental origin of mean racial-group differences is not supported. Although the Flynn effect does suggest that improving the environment, especially at the low end of the IQ distribution, can improve test scores, the cluster analysis shows that the *g* factor is independent of the Flynn effect. Instead, *g* is associated with inbreeding depression, for which there is no nongenetic explanation, which implies strongly that *g* is less amenable to environmental manipulation. These findings are consistent with an analysis of adoption data, which shows the IQ gains that result from being adopted into high SES homes do not produce a gain in *g* but only in non-*g* factors and in specificity of the various subtests. The adopted children's *g* factor scores reflected the SES level of their *biological* parents (Jensen, 1998a).

Dickens and Flynn (2001) replied to Rushton's (1999) cluster analysis with a more general statement of having resolved the paradox of how high heritabilities could go along with large secular increases in IQ. Their solution rests on the role

Table 4
Principal-Components Analysis and Varimax Rotation for Pearson Correlations of Inbreeding Depression Scores, Black–White Differences, g Loadings, and Gains Over Time on the Wechsler Intelligence Scales for Children (WISC) With Reliability Partialed Out

	Principal components			
	Unrotated loadings		Varimax rotated loadings	
Variable	I	II	1	2
Inbreeding depression scores from Japan (WISC–R)	0.31	0.61	0.26	**0.63**
Black–White differences from the U.S. (WISC–R)	0.29	0.70	0.23	**0.72**
WISC–R *g* loadings from the U.S.	−0.33	0.90	−0.40	**0.87**
WISC–III *g* loadings from the U.S.	−0.61	0.64	−0.66	**0.59**
U.S. gains 1 (WISC to WISC–R)	0.73	−0.20	**0.75**	−0.13
U.S. gains 2 (WISC–R to WISC–III)	0.81	0.40	**0.77**	0.47
German gains (WISC to WISC–R)	0.91	0.03	**0.91**	0.11
Austria gains (WISC to WISC–R)	0.87	0.00	**0.86**	0.07
Scotland gains (WISC to WISC–R)	0.97	0.08	**0.96**	0.17
% of total variance explained	48.60	25.49	48.44	25.65

Note. Values in bold define the factors. From "Secular Gains in IQ Not Related to the *g* Factor and Inbreeding Depression—Unlike Black–White Differences: A Reply to Flynn," by J. P. Rushton, 1999, *Personality and Individual Differences, 26,* p. 387. Copyright 1999 by Elsevier Science. Reprinted with permission.

of genotype–environment correlation. Recall from Section 5 that this occurs largely through the individual's genetic tendency to encounter, select, or create certain aspects of the environment in preference to alternatives. Genotype–environment correlation is part of the mechanism by which genetic proclivities become realized. Dickens and Flynn hypothesized that the positive feedback effects from even small initial environmental advantages stimulate mental development and lead to an even more favorable environment, stimulating yet more IQ development.

Dickens and Flynn's (2001) model, however, appears inconsistent with some empirical evidence. Gene–environment correlation cannot explain the mean Black–White group difference in IQ because it implies that Black groups, in comparison with White groups, become *increasingly* disadvantaged during the developmental period from early childhood to maturity. With increasing age there would be cumulative unfavorable effects on IQ for Black groups with respect to White groups. Yet national data (reviewed in Section 3) show that the size of the mean Black–White group difference remains approximately constant at 1 standard deviation and shows no significant change throughout the developmental period after about 3 years of age. The follow-up results of the Minnesota Transracial Adoption Study (Table 2), and the fact that the heritability for IQ generally increases from about 0.40 in *early* childhood to about 0.80 in *later* maturity (Figure 3), both contradict the Dickens–Flynn thesis. So too does the fact that both *g* estimates calculated

from East Indians in South Africa and genetic estimates calculated from the Japanese in Japan are able to predict the magnitude of Black–White differences in South Africa and in the United States (see Sections 4 and 5). Such robust generalization implies that the mean Black–White group difference in IQ is sufficiently persistent across cultures as to be unaffected by major changes in gene–environment correlations.

Dickens and Flynn (2001) provided no empirical evidence that gene–environment correlation constitutes either a large component of the phenotypic variance or that it increases with age (both of which are required by their model). They also did not provide any other direct empirical evidence. In addition, their models have been criticized for not taking the stability of IQ scores over time into account and for having too many free parameters (Loehlin, 2002; Rowe & Rodgers, 2002), to which Dickens and Flynn (2002) have replied. Because to date Dickens and Flynn have not given the high heritability of IQ any independent causal effect in explaining the mean Black–White group difference, their thesis is best placed in the culture-only camp.

Section 13: Evaluating the Culture-Only and the Hereditarian Research Programs

Table 5 summarizes the 30-plus-years of research on Black–White IQ differences carried out since Jensen's (1969) *Harvard Educational Review* article. It compares and contrasts the predictions of the hereditarian and the culture-only theories against the existing data reviewed in Sections 3 through 12, to which we then assigned "scores." We assigned the highest score (++) when, in our opinion, the results confirmed a novel prediction first derived from theory that was then tested and confirmed. We awarded the next highest score (+) when the results were consistent with theory but not predicted from it. We gave a neutral score (0) when the results could not be predicted from theory so that it could be either confirmed or disconfirmed. We assigned a negative score (–) when the predicted results were disconfirmed. Because some diacritical tests have two components, the maximum possible support for either research program would be a score of $12 \times 2 = 24$; maximum disconfirmation would be a score of –24. Naturally these scores reflect our particular evaluation of how well the two competing theories predict and explain the evidence. We acknowledge that others might see things differently, and we invite them to assign their scores. Our purpose is to advance the debate.

Our evaluation of the evidence supports a cumulative score of 17 for the hereditarian model and –7 for the culture-only model. We therefore suggest that the hypothesis of *some* genetic component in the mean Black–White group difference in IQ should be considered "provisionally true." Naturally, we do not expect everyone to agree with this assessment. Our own perspective is obviously hereditarian (Jensen, 1998b; Rushton, 2000). Those working from a different perspective may arrive at alternative tallies or add new dimensions to be tallied that we have overlooked. Before discussing our conclusion, we consider in more detail the data on each of the categories in Table 5.

Table 5
Comparison Chart for Evaluating the Hereditarian (50% Genetic) and Culture-Only (0% Genetic) Research Programs

Section: Prediction	Hereditarian model	Score	Culture-only model	Score
3: Mean Black–White–East Asian IQ differences found worldwide.	Predicted by hereditarian model.	(++)	Lower Black scores due to cultural deprivation, but higher East Asian scores anomalous.	(−)
4: Mean Black–White differences greater on *g* loaded cognitive tests than on culturally loaded cognitive tests.	Novel prediction from theory confirmed.	(++)	Predictions contradicted.	(−)
5: Mean Black–White differences in IQ greater on more heritable subtests than on culturally malleable subtests.	Novel prediction from theory confirmed.	(++)	Predictions contradicted.	(−)
6: Mean Black–White–East Asian IQ differences associated with mean differences in brain size.	Explains Black–White IQ differences.	(++)	Lower Black size due to nutritional deprivation, but East Asian size anomalous.	(−)
7: Mean Black–White–East Asian IQ differences remain (in all directions) following transracial adoption, despite early deprivation in East Asians.	Genetic effects stronger than common environment effects.	(++)	Prediction fails because Black–White IQ differences do not disappear. Cannot explain why East Asians average higher than Whites.	(−)
8: Mean Black–White differences in IQ and brain size reflected in studies of racial admixture. Children wrongly believed by adoptive parents to have two Black biological parents had the same IQs as children correctly classified as of mixed race.	Predicted by genetic theory.	(+)	Social discrimination via skin color does not explain racial misclassification nor the pattern of mean brain size differences.	(0)
9: Mean Black–White IQ differences show regression toward predicted race means. Best evidence comes from well-controlled sibling studies.	Novel prediction confirmed.	(++)	Parent–child regression predicted, but sibling data anomalous.	(0)

Table 5 (*continued*)

Section: Prediction	Hereditarian model	Score	Culture-only model	Score
10: Mean Black–White–East Asian IQ differences are paralleled by a race–behavior matrix of 60 life-history traits.	Series of novel predictions based on an evolutionary theory of racial origins confirmed.	(++)	Anomalous.	(−)
11: Mean Black–White–East Asian differences mesh with what is known about human evolution.	Series of novel predictions based on evolutionary theory confirmed.	(++)	Anomalous.	(−)
12: Mean Black-White-East Asian differences not fully explained by any model, not by a 100% culture-only model, by a 100% genetic-only model, or, with precision, even by the 50% genetic-50% environmental model.	The model of 50% genes-50% environment may need to be revised, perhaps to 80% genes-20% environment.	(0)	Model as currently stated fails. Identifying some large Gene × Environment interaction might make it viable.	(0)
Total		17		−7

Note. Key for scores: ++ = strong support: results confirmed a novel prediction first derived from theory and then tested; + = mild support: results consistent with theory but not predicted from it; 0 = neutral: results could not be predicted from theory so that it could be either confirmed or disconfirmed; − = rejection: results predicted from theory disconfirmed. Because some diacritical tests have two elements, the maximum possible support would be a score of $12 \times 2 = 24$ and maximum disconfirmation would be a score of −24.

Mean Race–IQ Differences Are Found Worldwide (Section 3)

The mean Black–White IQ difference in the United States of 85 versus 100 can be, and has been, explained both by the hereditarian model (in terms of some genetic difference) and by the culture-only model (in terms of nutrition, poverty, SES, family structure, schooling, racism, and the legacy of slavery). Hence, initially we were inclined to give both the hereditarian model and the culture-only model a score of (+). The hereditarian model, however, also predicted that the same pattern would be found worldwide, with lower scores for sub-Saharan Africa than for Black Americans, and that the differences would also be found on culture-fair tests and on reaction time tasks that measure the speed and efficiency with which the brain processes information (and which all children can perform in less than 1 s). These predictions were confirmed. The culture-only hypothesis is disconfirmed by the differences on culture-fair and reaction time tests. Nor can the culture-only model easily explain why the East Asian average IQ of 106 is *higher* than the average White IQ, including on these same speed-of-processing tasks. Within the United States, the mean Black–White group difference in IQ has not changed significantly over the past 100 years despite significant improvements in the conditions of Black Americans. The same magnitude of difference is observed as early as age 2½ years. Our score for Section 3: hereditarian model (+); culture-only model (–).

Race–IQ Differences Are Most Pronounced on the More g-Loaded Components of Tests and Least So on the Most Culturally Loaded Items (Section 4)

The hereditarian model made the novel prediction that the mean Black–White group difference in IQ is not the result of idiosyncratic cultural peculiarities in this or that test but would be more pronounced on highly *g*-loaded tests. Because the prediction was confirmed, it counts as evidence for the hereditarian position while also contradicting the prediction from the culture-only model that the differences are due to culturally loaded tests. In South Africa, *g* loadings calculated on East Indians predicted mean Black–White group differences, showing substantial cross-cultural generalizability. A test's *g* loading is the best predictor, not just of its correlation with scholastic and workplace performance, but also of its correlation with reaction time measures, heritability coefficients determined from twin studies, inbreeding depression scores calculated in children of cousin-marriages, and neurological variables such as brain size, brain evoked potentials, brain pH levels, brain glucose metabolism, and nerve conduction velocity. Thus, we conclude the evidence reviewed in Section 4 strongly supports the hereditarian model (++) and argues against the culture-only model (–).

Race–IQ Differences Are Most Pronounced on the More Heritable Components of Tests With Little or No Evidence of Race-Specific Developmental Processes (Section 5)

Research based on this novel prediction from the hereditarian model established that (a) the heritability of IQ among Black groups is around 0.50, not significantly different from that found in White groups; (b) there is no evidence

of the effect of any special minority-specific developmental process resulting from the legacy of slavery or of White racism in large sets of archival correlation matrices between background variables and outcome measures, or on relations among subtests; (c) IQ subtests with higher heritabilities predict mean Black–White group differences better than do subtests with lower heritabilities; and (d) the shared environment type of variables usually proposed to explain group differences (e.g., differences in income, schools) decrease in effect size with age. Cross-cultural generality was demonstrated by the fact that the magnitude of inbreeding depression scores on various subtests calculated on the Japanese in Japan predicted the magnitude of Black–White differences in the United States. Because the empirical results confirmed a novel prediction from the hereditarian model (++) but disconfirmed several predictions from culture-only theory (–), we scored Section 5: hereditarian model (++); culture-only model (–).

Mean Race–IQ Differences Are Associated With Mean Brain Size Differences (Section 6)

Overall, MRI studies show that brain size is related to IQ differences within race. Moreover, the three-way pattern of group differences in average brain size is detectable at birth. By adulthood, East Asians average 1 cubic inch more cranial capacity than Whites, and Whites average 5 cubic inches more cranial capacity than Blacks. These findings on group differences in average brain size have been replicated using MRI, endocranial volume from empty skulls, wet brain weight at autopsy, and external head size measures. They were acknowledged by Ulric Neisser, Chair of the APA's Task Force on intelligence, who noted that, with respect to "racial differences in the mean measured sizes of skulls and brains (with East Asians having the largest, followed by Whites and then Blacks) . . . there is indeed a small overall trend" (Neisser, 1997, p. 80). The hereditarian model explains these in terms of genetic differences. The culture-only position can explain them in terms of nutrition, SES, or early cognitive stimulation. Adding the East Asian data, however, literally "changes the shape of the table." The hereditarian model posits that if East Asians average higher IQs than do Whites, then they must also average larger brains than Whites, and that perhaps both the higher IQ and the larger brain are most parsimoniously explained in terms of the natural selection experienced in colder climates during human evolution (++). The culture-only position has yet to explain both the higher IQ *and* the larger brain size of East Asians, given that these groups have also been subjected to prejudice in White societies or severe malnutrition in their homelands. We scored Section 6: hereditarian model (++); culture-only model (–).

Mean Race Differences in IQ Remain Following Transracial Adoption (Section 7)

Transracial adoption studies provide one of the best methods for resolving the question of group differences in mean IQ. The above-average IQ scores of Black adoptees at age 7 confirmed the culture-only predictions. The results of the follow-up testing at age 17 were more ambiguous. The hereditarian model predicted that when the longitudinal study was carried out, the Black–White difference would emerge (based on the increasing size of the genetic effect on IQ

with age), and this is one interpretation of the data, though a culture-only interpretation is also plausible. However, support for the hereditarian model again comes from adding the East Asian data to the mix. Korean and Vietnamese children adopted into White homes, even though as babies many had been hospitalized for malnutrition, nonetheless grew to have IQs 10 or more points higher than their adoptive national norms. The culture-only model cannot explain that finding. Further, it argues against the culture-only hypothesis that the high performance of East Asian children is due to "trying harder" or other cultural values emphasized by East Asian families. Our score for Section 7: hereditarian model (++); culture-only model (–).

Studies of Racial Admixture Reflect Mean Black–White IQ Differences (Section 8)

Both the hereditarian and the culture-only model can explain why groups of lighter skinned African Americans and the (also lighter skinned) mixed-race "Coloreds" of South Africa have average IQs between those of (for the most part) unmixed groups of Blacks and Whites. Both models can also explain the fact that the degree of White admixture is correlated with brain weight at autopsy. The culture-only position does so in terms of societal discrimination based on skin color as well as its possible cascading effects on nutrition and health (+); the hereditarian model does so in terms of the hypothesized genetic difference in average IQ and its correlations with race and skin color (+). Some evidence against the culture-only position comes from studying the misclassified adoptees in the Minnesota Transracial Adoption Study (–). The expectancy effects hypothesis, that adoptive parents' beliefs about their child's racial background influence the child's intellectual development, is not supported by the finding that the mean IQ score for 12 children wrongly believed by their adoptive parents to have had two Black biological parents was about the same as that of the 56 children correctly believed by their adoptive parents to have had one Black and one White biological parent. While the number of children is small, this conclusion is supported by Rowe's study in which 119 mixed-race children were selected as "looking African American" but their IQ scores also turned out to be intermediate. Our score for Section 8: hereditarian model (+); culture-only model (0).

IQs Show Regression Toward Predicted Racial Means (Section 9)

The phenomenon of regression to the mean is predicted from basic genetic theory and has been documented for a number of physical traits in humans and in other species. The hereditarian model applied this reasoning to IQ studies to make a novel prediction about the amount of regression across the *whole* IQ distribution and *various* degrees of kinship. The results showed that the children of Black parents of IQ 115 regressed toward the Black average IQ of 85, whereas children of White parents of IQ 115 regressed toward the White average IQ of 100. The converse occurred at the low end of the scale. Even stronger support for the hereditarian position came from sibling data. The regression lines for both Blacks and for Whites showed no significant departure from linearity throughout the range of IQ from 50 to 150. A failure of this prediction would have argued against the hereditarian model but would have been neutral for the culture-only model.

The predictions from the hereditarian model were tested and confirmed. The culture-only theory must argue that environmental effects or chance variation mimics the predicted genetic effects. We scored Section 9: hereditarian model (++); culture-only model (0).

Mean Race–IQ Differences Are Paralleled by a Matrix of Other Traits and Behaviors (Section 10)

A suite of over 60 life-history variables, including rate of two-egg twinning, speed of maturation and longevity, personality and temperament, family stability and crime, sexual behavior and fertility, as well as intelligence and brain size, have been identified on which East Asian and African groups consistently average at the two ends of a continuum, with European groups intermediate, regardless of where they presently live. This race–behavior matrix constitutes a series of novel predictions derived from an evolutionary theory of the origin of races that were tested and confirmed. The culture-only model has only partially addressed this race–behavior matrix, with (sometimes contradictory) supplementary hypotheses. Our score for Section 10: hereditarian model (++); culture-only model (–).

Mean Race–IQ Differences and Human Evolution (Section 11)

One theory of human evolution argues that the farther north the ancestral human populations migrated out of Africa, about 100,000 years ago, the more they encountered the cognitively demanding problems of gathering and storing food, gaining shelter, making clothes, and raising children successfully during prolonged winters. (This is not the only theory of human evolution, nor do all who endorse it concur with our interpretation.) Ecological pressures selected for larger brains, slower rates of maturation, lower levels of sex hormone, and all the other life-history characteristics. From this perspective, the data from both human genetics and human evolution mesh with the race–behavior matrix (++). Genetic–evolutionary theory acknowledges factors such as East Asian family strength or African poverty, but as effects rather than causes. The consistency of the pattern of traits in Table 3 also supports the argument, as do genetic analyses, against the view that race is only a social construction based on a few salient traits such as skin color (Crow, 2002; Risch et al., 2002). Our score for Section 11: hereditarian model (++); culture-only model (–).

Culture-Only Hypotheses Fail to Account for Mean Race–IQ Differences (Section 12)

This section reviewed a number of well-known culture-only hypotheses for the mean Black–White group difference in IQ. The most widely accepted is that they are due to differences in SES. Adjusting for SES, however, only reduces the mean Black–White IQ difference by about one third. Other culture-only hypotheses, such as the effects of segregation, bias in tests, or the consequences of being a minority in a White society are not supported by our review of the evidence. The Head Start program in the United States has produced some modest gains in increasing school retention and graduation rates among White groups, though not among Black groups (Currie & Thomas, 1995). Neither the narrowing of the

Black–White social conditions nor the Flynn effect (i.e., the secular rise in IQ) has narrowed the Black–White IQ gap. However, the Flynn effect, based on increases in nutrition, health care, and intellectual stimulation, does appear to support the culture-only model (+), but it is neutral (hence 0) for the hereditarian model because large environmental effects (up to 50% of the variance) are compatible with large genetic effects (also up to 50% of the variance). The finding in Table 4 that the secular increase did not cluster with *g* and its biological correlates, however, may support the hereditarian model. Because the real area of conflict is the cause of the mean racial group difference in cognitive ability, the hereditarian hypothesis is not disproven by the Flynn effect, whatever its cause(s). Overall, on Section 12, we gave the hereditarian model a score of (0) and the culture-only model (0).

Section 14: Progressive Research Leads to Provisional Truth

Our conclusion, that the Black–White IQ difference is partly heritable, accords with previous analytic reviews of this literature. Loehlin et al. (1975) concluded that Black–White IQ differences "probably" reflected "genetic differences among the groups" (p. 238). P. E. Vernon (1979) tabulated 30 main topics, each scored on a 4-point scale, and concluded that "although the total number of items favoring genetic influences (G and G?) is roughly balanced by the number of environmental points (E and E?), more of the highly convincing items are G rather than E" (p. 319). The survey of over 1,000 experts in behavioral genetics and psychometrics by Snyderman and Rothman (1987) also found that a plurality believed the Black–White IQ difference "to be a product of both genetic and environmental variation" (p. 141). However, there are also notable statements to the contrary. The APA Task Force on intelligence, for example, concluded "[t]here is certainly no support for a genetic interpretation" (Neisser et al., 1996, p. 97). Likewise, Nisbett (1998) reached the conclusion that "the most relevant studies provide no evidence for the genetic superiority of either race" (p. 101).

In our opinion, the present review, similar to those of Loehlin et al. (1975) and Vernon (1979) earlier, should be given greater weight because they surveyed a greater range of evidence. Examining all the documentation allows a greater chance of finding accurate explanations than does selecting a few items from the whole. The 10 categories of predictions reviewed in Table 5 were derived from the "hard core" assumptions of the two competing research programs, each of which tries to explain the Black–White IQ difference (see Section 2). Based on Lakatos's (1970, 1978) criteria for evaluating research programs, and a philosophy of science methodology that evaluates rival theories by generating multiple strong inferences and assessing the preponderance and the consilience of many lines of evidence, we believe the hereditarian theory has satisfied the criteria for a "progressive" research program, whereas the culture-only program has not. Both have drawn implications to make numerous, testable, novel predictions, but we found the hereditarian predictions were mostly confirmed, whereas those from culture-only theory mostly were not.

Some attempts to salvage the culture-only theory have been problematic. Suggestions that "White flight" undermined desegregation and busing programs so that "true" integration has not yet been properly tried, that Head Start programs

still have not been fully funded, that culture-fair tests were not *really* fair, that the sensory deprivation and race-of-examiner hypotheses were of only minor significance, or that being Black in America is really a matter of caste, not class or race, fall short of being strong inference. Urbach (1974b, p. 237) dismissed such claims, and that the mean Black–White IQ difference is a system problem, and that caste and class differences are complex, subtle, and attitudinal rather than grossly econometric, as "pseudo-scientific maneuvers," noting that "it is *intellectually improper to obscure facts by continually retreating behind the trivial truth that the world is complex* [italics added]".

We believe the burden of proof must shift to those who argue for a 100% culture-only position. For example, they need to address why, if important minority-specific developmental processes such as stereotype threat (Steele, 1997) and racial stigma (Loury, 2002) exert such a powerful influence on school achievement, the correlation matrices representing developmental processes can be so similar across ethnic and racial groups (Section 5). They need to explain why, if gene–environment interactions are as widespread and difficult to disentangle as often claimed (e.g., Block, 1995), identical twins reared apart grow to be so similar (Bouchard, 1996; Bouchard & Loehlin, 2001). Some culture-only hypotheses are too ambiguous to be tested.

A conundrum for theorists of all persuasions, however, is that there is too little evidence of any environmental effects. The hereditarian model of Black–White IQ differences proposed in Section 2 (50% genetic *and* 50% environmental), far from precluding environmental factors, requires they be found. Although evidence in Sections 3 to 11 provided strong support for the genetic component of the model, evidence from Section 12 was unable to identify the environmental component. On the basis of the present evidence, perhaps the genetic component must be given greater weight and the environmental component correspondingly reduced. In fact, Jensen's (1998b, p. 443) latest statement of the hereditarian model, termed the *default hypothesis*, is that genetic and cultural factors carry the exact same weight in causing the mean Black–White difference in IQ as they do in causing individual differences in IQ, about 80% genetic–20% environmental by adulthood.

One current challenge, therefore, is to identify significant sources of nongenetic variance. It is hoped that the recent models of gene–culture correlation put forward by Dickens and Flynn (2001, 2002) to explain the paradox of large heritability estimates and large environmental effects mark a new trend in the race–IQ debate. Their models accept the empirical reality of both genetic and environmental influences. Within Whites, most of the nongenetic variance appears to result from a small number of random adverse effects such as prenatal problems, complications in the birth process, maternal health, and childhood disease and trauma (Jensen, 1997).

Another challenge is to explain the pattern of covariant traits shown in Table 3 and described in Section 10. Several life-history theorists working from an evolutionary perspective have postulated that there is some latent dimension on which "unpredictable environments" and "scarcity of resources" can move individuals up or down because traits need to be coherent and harmonized, rather than having some go to one pole while others go to the other (see Rushton, 2000, pp. 252–255, 271–273, for discussion). Most life-history explanations that have

focused on behavioral rather than biological traits have typically avoided the race differences and have hypothesized postnatal events such as father absence, parental divorce, and sexual and physical abuse as causes rather than effects (e.g., Belsky, Steinberg, & Draper, 1991; Chisholm, 1999). However, the fact that brain size differences show up at birth, as do those in gestation time and speed of physical maturation, implies some prenatal and biological mediation.

Writing from a life-history perspective, Mealey (1990) suggested that the wider pattern of racial-group differences described in Section 10 was "interesting and worth pursuing, but . . . may be environmentally contingent rather than genetic" (p. 387). Nyborg (1994, pp. 146–149) hypothesized that an estradial-testosterone hormonal trade-off existed, with Blacks averaging the most testosterone, East Asians the least, and Whites in between. Eysenck (1991) proposed a nutritional deficiency hypothesis. Lynn (1990; Lynn & Vanhanen, 2002) also proposed a crucial role for nutrition, finding that the heavier twin at birth typically has a larger head and brain size, as well as a higher IQ in adolescence, and that infants fed breast milk typically average higher IQs (8.3 points) at age 8 than those fed formulas. Masters (1997, p. 142) put forward a "neurotoxity hypothesis," in which race differences in pre- and postnatal exposure to metal pollutants interact with stress, poor diet, alcohol, and drugs. He pointed out that breast-feeding reduces the infants' exposure to metal pollutants while providing infants the long chains of proteins necessary for brain development. Mothers from low socioeconomic groups typically do not breast-feed their infants, and Black mothers are only one third as likely to breast-feed their infants as White mothers (see also Jensen, 1998b, pp. 506–508).

Section 15: Implications for Public Policy

It is a widely accepted fact of behavioral science that there is great variability within each racial group and there is an ethical consensus that we treat people as individuals. Although no specific policies necessarily follow from knowing about the causes of group differences, they may serve as guides to action on some issues. The conclusion reached in Sections 13 and 14—that about 50% of the variance in mean Black–White group differences in IQ is due to heredity—is compatible with a wide range of recommendations, from programs for the disadvantaged and laissez-faire approaches to selection and opportunity grouping in certain educational and vocational situations.

In *The Bell Curve,* Herrnstein and Murray (1994) offered some specific policy recommendations based on their conclusions about genetic variation and IQ, which are generally concordant with political conservatism, such as scaling back affirmative action, reducing the intrusiveness of government, and returning to individualism. Most political conservatives, however, support these recommendations, no matter how the nature–nurture question is "resolved," an argument with which Murray agreed (Miele, 1995). Arthur Jensen, also writing from the hereditarian perspective, recently opined that giving primacy to individual rights maximizes fairness, which he pragmatically defined as the ability of each individual to reach his or her full potential (Miele, 2002). He therefore argued for a restructuring of the educational system by tailoring methods to fit the *individual*,

and letting the *group* outcomes become what they may, rather than allowing claims of differential performance to justify group rights over individual rights.

A multifaceted approach to the policy implications of the hereditarian conclusion about mean Black–White differences is required. Many policy preferences are not at all affected by our scientific conclusion. Granting equal rights under the law and whether to provide social welfare, for example, are based more on moral and political philosophy than on research findings. Still other policy issues (e.g., affirmative action, the value of diversity) might merit reconsideration based on the degree to which heredity as opposed to culture turns out to be the causal agent. Finally, certain policies do follow from our conclusions, such as the need for greater equality and neutrality in the treatment of the culture-only and genetic hypothesis within both the scientific and policy arenas, and in researching and ameliorating the biological basis of group differences in IQ, health, and so on.

Two fundamentally different, policy-organizing models typically used to explain why racial groups differ in average rate of socially valued outcomes should be examined, specifically, with reference to the issues of race relations, educational and psychological testing, health, and conflicting worldviews about the nature of human nature.

Discrimination or Distribution?

Herrnstein (1990) termed the two fundamentally different models put forth to explain why racial groups differ in their average rate of socially valued outcomes the distributional model and the discrimination model. Each may be partially correct. The *discrimination model* focuses on social and institutional practices that discriminate against members of one group (or favor members of another), thus tilting the playing field. It assumes that in the absence of discrimination, outcomes should be about equal for all populations; thus evidence of differential performance in itself constitutes evidence of discrimination. Factors hypothesized under this model that cause mean race differences include relative poverty, anti-Black bias, a lack of access to legitimate channels of upward mobility, and dysfunctional family organization growing out of the legacy of slavery (see Sections 12 and 13).

The discrimination model has also been used to explain the *overrepresentation* of some groups in valued outcomes. Blacks are said to excel in sports such as boxing, basketball, track and field, and football because other channels of upward mobility are closed to them (Hoberman, 1997). As early as the 1920s, sociologists explained the underrepresentation of East Asians in U.S. crime statistics as being due to the East Asian "ghetto." This self-imposed segregation was seen as a response to external prejudice, which protected its members from the disruptive tendencies of the outside society.

The *distributional model,* on the other hand, explains the overlapping of the racial groups and their differing averages in terms of their mean group characteristics—for example, the mean differences in heritable IQ and possibly other traits too. However, it could also fit Sowell's (1994) theory of socialization through subtle cultural traditions, or Loury's (2002) theory of racial stigma, which postulates a unique type of gene–culture correlation in which people react to others on the basis of physical appearance. Other factors hypothesized to underlie a distributional model include deep-rooted cultural values and family structures

endemic to certain populations, as well as biological variables such as body type, hormonal levels, and personality and temperament. Thus according to the distributional model, population differences are expected to occur and to do so globally.

The research supporting the role of heredity in human behavior implies that the distributional model is more correct than the discrimination model. It explains some of the mean Black–White group difference in IQ-related outcomes in terms of the differential distribution of the genes for general mental ability. For example, IQ is a significant predictor of such socially disadvantageous outcomes as dropping out of high school, being unemployed, being divorced within 5 years of marriage, having an illegitimate child, living in poverty, being on welfare, and incarceration. In today's technological society, everyday life itself is a type of IQ test (see R. A. Gordon, 1997; Gottfredson, 1997; Herrnstein & Murray, 1994). On all of the above measures, the group means favor Whites over Blacks. Of course, this does not deny that many other attributes are also important for success in life.

Race Relations

Some have suggested that we cannot expect members of ethnic groups to simply accept the genetic component in the mean-group differences in IQ and other traits. Yet, with regard to individuals within families, we do acknowledge that some siblings are more intelligent, more athletic, more physically attractive, or more socially charming than others. We also accept that some families are genetically more gifted in certain areas than other families. We should, therefore, by extension, be able to generalize to all the members of the human family. If viewed against the backdrop that group differences are simply aggregated individual differences, the former may be easier to accept than has hitherto been thought.

Although the hereditarian model may not specify particular policies, it does argue against the feasibility of some programs based on the discrimination model. One very socially significant example is that a demonstration of differential racial performance (good or bad) could not, by itself, be offered as proof of racial discrimination because, as the evidence reviewed in this article demonstrated, genetic factors play a role in producing these differences. Rather, the burden would be on the plaintiffs to prove that the defendants had discriminated on the basis of race and not educational or vocational performance associated with race.

However, the view that one segment of the population is largely to blame for the problems of another segment can be even more harmful to racial harmony, by first producing demands for compensation and thereby inviting a backlash. Equating group disparities in success with racism on the part of the more successful group guarantees mutual resentment. As overt discrimination fades, still large racial disparities in success lead Blacks to conclude that White racism is not only pervasive but also insidious because it is so unobservable and "unconscious." Whites resent that nonfalsifiable accusation and the demands to compensate Blacks for harm they do not believe they caused. Misplaced blame can also endanger institutions. Objective standards of merit that yield racially uneven outcomes when evenly applied, whether in college admissions, hiring, or day-to-day operations of the legal system, have increasingly come under attack—for

example, by critical legal theorists as inherently pro-White and hence illegitimate. Although the distributional model does not rule out affirmative action or compensation-type initiatives, it does reduce the impact of arguments in their favor based on an exclusive adherence to the discrimination model (Gottfredson, 2000; Levin, 1997).

Educational, Vocational, and Psychological Testing

Because the means for Blacks and Hispanics are lower on tests of academic and vocational achievement, such as the SAT, the General Aptitude Test Battery, and the ASVAB, than those for Whites and East Asians, some have claimed the tests are racially biased. Yet the evidence reviewed and the distributional model predict that such differences will occur worldwide (see Section 3). This is supported by the fact that these tests have about equal predictive validity for all groups who speak the same language and have been schooled in the culture of the test.

Ethnic disparities in cognitive performance are not just "an American dilemma" (Myrdal, 1944) but are found around the world. In India, members of the higher castes obtain higher mean scores and examination marks than do those of the lower castes. In Malaysia, members of the Chinese and East Indian racial minorities have higher mean scores than does the majority Malay population. In South Africa, members of the White, East Indian, and Colored population groups obtain higher mean scores than does the indigenous Black African majority (Klitgaard, 1986; Lynn & Vanhanen, 2002; Nell, 2000). These facts present a challenge for policymakers attempting to design educational, vocational, and military systems.

The apparent failure of equal opportunity programs to enable all groups in society to perform equally scholastically or even to narrow the gap in the test scores used in selection for higher education, jobs, and the Armed Forces has caused some to disparage psychological testing. Jensen (2000) proposed emphasizing *face validity* in test construction by making item content more obviously relevant to the purpose of selection and thereby improve the public's perception of the utility of tests. An example would be having all the items in a test used to select police personnel involve a typical crime scenario rather than abstract analogies (though a correct solution would still require logical reasoning).

More generally, there is a need to educate the public about the true nature of individual and group differences, genetics, and evolutionary biology. Ultimately, the public must accept the pragmatic reality that some groups will be overrepresented and other groups underrepresented in various socially valued outcomes. Organizations such as the APA could play a critical role in changing the zeitgeist. To do so will not be easy, for it requires overcoming deeply ingrained biases that operate at several levels of the APA (Redding, 2001). The standard models of social science from the 1930s to the present have assumed a tabula rasa perspective that precludes any analysis of hereditary group differences.

The expectation of average group differences includes the statistical certainty that it does not apply to all individuals, a point that can easily be overlooked. Thus, any part of a general program of education must include distributional statistics so that people also learn not to stereotype or *over*generalize. This may

not be as difficult a task as might be supposed. Even kindergarten children are capable of learning that although boys are typically taller than girls, many girls are taller than the average boy.

Health, Medical Genetics, and Pharmaco-Anthropology

The distributional model also relates to questions raised about bias in the health system based on ethnic differences in rates of certain diseases (Risch et al., 2002). From an epidemiological perspective, failure to take ethnicity into account can confound a study, particularly if the disease in question is more common in one ethnic group than another. The incidence of hypertension, prostate cancer, and other diseases is higher in Blacks than Whites. To cite a specific example, the infant mortality rate of American Blacks is twice that of American Whites, and this difference remains significant after controlling for SES in each race.

Despite overlap, drug and food effects often differ sufficiently to test for them independently. One well-known example of the harm done by ignoring group differences involves lactose intolerance. The ability of adults to digest milk easily is largely limited to Caucasoids, and a lack of this knowledge may have increased mortality among the needy in Third World countries that were inadvertently provided with milk products to alleviate hunger.

Some physicians are becomingly increasingly concerned that assuming Blacks are identical to Whites neglects the problems faced by Blacks. For example, 30% of the people who have kidney failure and undergo dialysis are Black, but estimates are that fewer than 10% of organ donors are Black. There is also evidence that Blacks fare better when given organs donated by other Blacks. Another example is that it now appears that genetics may underlie the increased rates and levels of hypertension in Blacks. Black males tend to have higher blood pressure and increased rates of cardiovascular disease, including strokes, than White males, but the cause is widely debated. Some evidence suggests that Black males experience a faster heart rate when performing moderate exercise, although the pulse rates of the Black and White males when resting show no significant difference. Blacks have a higher incidence of prostate cancer than Whites, who in turn have a higher rate than East Asians, for which the underlying cause may be the mean group differences in testosterone level (Polednak, 1989; Risch et al., 2002).

Conflicting Worldviews

A prevailing worldview throughout history has been that economic, cultural, and other environmental forces are the preeminent causes of group and individual behavior. Modern social science has typically taken this perspective and promoted the idea that all babies are born more or less equally endowed in intelligence and learning ability. It followed therefore that inequalities were the result of social, economic, and political forces. This worldview generated many strategies for intervention in the home, the workplace, the mass media, the criminal justice system, and even the entire social–economic system. Some have been effective and are almost universally accepted, whereas others have failed and produced only shattered expectations, resentment, and interethnic hostility.

The major policy implication of the research reviewed here is that adopting an

evolutionary–genetic outlook does not undermine our dedication to democratic ideals. As E. O. Wilson (1978) aptly noted: "We are not compelled to believe in biological uniformity in order to affirm freedom and dignity" (p. 52). He went on to quote the sociologist Bressler (1968): "An ideology that tacitly appeals to biological equality as a condition for human emancipation corrupts the idea of freedom. Moreover, it encourages decent men to tremble at the prospect of 'inconvenient' findings that may emerge in future scientific research" (E. O. Wilson, 1978, p. 52). Denial of any genetic component in human variation, including between groups, is not only poor science, it is likely to be injurious both to unique individuals and to the complex structure of societies.

References

Bartley, A. L., Jones, D. W., & Weinberger, D. R. (1997). Genetic variability of human brain size and cortical gyral patterns. *Brain, 120,* 257–269.

Bayley, N. (1965). Comparisons of mental and motor test scores for ages 1–15 months by sex, birth order, race, and geographic location, and education of parents. *Child Development, 36,* 379–411.

Beals, K. L., Smith, C. L., & Dodd, S. M. (1984). Brain size, cranial morphology, climate, and time machines. *Current Anthropology, 25,* 301–330.

Bean, R. B. (1906). Some racial peculiarities of the Negro brain. *American Journal of Anatomy, 5,* 353–432.

Belsky, J., Steinberg, L., & Draper, P. (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. *Child Development, 62,* 647–670.

Berry, J. W. (1966). Temne and Eskimo perceptual skills. *International Journal of Psychology, 1,* 207–222.

Block, N. (1995). How heritability misleads about race. *Cognition, 56,* 99–128.

Bouchard, T. J., Jr. (1996). Behavior genetic studies of intelligence, yesterday and today: The long journey from plausibility to proof (The Galton Lecture). *Journal of Biosocial Science, 28,* 527–555.

Bouchard, T. J., Jr., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics, 31,* 243–273.

Brazelton, T. B., & Freedman, D. G. (1971). The Cambridge neonatal scales. In G. B. A. Stoelinga & J. J. van der Werff ten Bosch (Eds.), *Normal and abnormal development of brain and behavior* (pp. 104–132). Leiden, the Netherlands: Leiden University Press.

Breslau, N., Chilcoat, H. D., Susser, E. S. Matte, T., Liang, K.-Y., & Peterson, E. L. (2001). Stability and change in children's intelligence quotient scores: A comparison of two socioeconomically disparate communities. *American Journal of Epidemiology, 154,* 711–717.

Bressler, M. (1968). Sociobiology, biology and ideology. In D. Glass (Ed.), *Genetics* (pp. 178–210). New York: Rockefeller University Press.

Broca, P. (1873). Sur les crânes de la caverne de l'Homme Mort (Loere) [Cranial size in the deceased]. *Revue d'Anthropologie, 2,* 1–53.

Broman, S. H., Nichols, P. L., & Kennedy, W. (1975). *Preschool IQ: Prenatal and early developmental correlates.* Hillsdale, NJ: Erlbaum.

Broman, S. H., Nichols, P. L., Shaughnessy, P., & Kennedy, W. (1987). *Retardation in young children*. Hillsdale, NJ: Erlbaum.

Brown v. Board of Education of Topeka, 347 U.S. 483 (1954).

Bulmer, M. G. (1970). *The biology of twinning in man*. Oxford, England: Clarendon Press.

Carretta, T. R., & Ree, M. J. (1995). Near identity of cognitive structure in sex and ethnic groups. *Personality and Individual Differences, 19,* 149–155.

Cavalli-Sforza, L. L. (2000). *Genes, peoples, and languages*. Berkeley: University of California Press.

Cavalli-Sforza, L. L., Menozzi, P., & Piazza, A. (1994). *The history and geography of human genes*. Princeton, NJ: Princeton University Press.

Chakraborty, R., Kamboh, M. I., Nwankwo, M., & Ferrell, R. E. (1992). Caucasian genes in American Blacks. *American Journal of Human Genetics, 50,* 145–155.

Chan, J., & Lynn, R. (1989). The intelligence of six-year-olds in Hong Kong. *Journal of Biosocial Science, 21,* 461–464.

Chisholm, J. S. (1999). *Death, hope, and sex: Steps to an evolutionary ecology of mind and morality*. Cambridge, England: Cambridge University Press.

Clark, E. A., & Hanisee, J. (1982). Intellectual and adaptive performance of Asian children in adoptive American settings. *Developmental Psychology, 18,* 595–599.

Cohn, S. J., Cohn, C. M. G., & Jensen, A. R. (1988). Myopia and intelligence: A pleiotropic relationship? *Human Genetics, 80,* 53–58.

Coleman, J. S. (1990–1991). The Sidney Hook Memorial Award Address: On the self-suppression of academic freedom. *Academic Questions, 4,* 17–22.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Department of Health, Education, and Welfare.

Cooley, C. H. (1897). Genius, fame, and the comparison of races. *Annals of the American Academy of Political and Social Science, 9,* 317–358.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12,* 671–684.

Crow, J. F. (2002, Winter). Unequal by nature: A geneticist's perspective on human differences. *Daedalus, 131*(1), 81–88.

Currie, J., & Thomas, D. (1995). Does Head Start make a difference? *American Economic Review, 85,* 341–364.

Darwin, C. (1859). *The origin of species.* London: Murray.

Darwin, C. (1871). *The descent of man*. London: Murray.

Davenport, R. K. (1946). Implications of military selection and classification in relation to universal military training. *Journal of Negro Education, 15,* 585–594.

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain.* Oxford, England: Oxford University Press.

Degler, C. N. (1991). *In search of human nature*. New York: Oxford University Press.

Devlin, B., Feinberg, S. E., Resnick, D. P., & Roeder, K. (1997). (Eds.). *Intelligence, genes, and success: Scientists respond to The Bell Curve.* New York: Springer-Verlag.

Dickens, W. T., & Flynn, J. R (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108,* 346–369.

Dickens, W. T., & Flynn, J. R. (2002). The IQ paradox is still resolved: Reply to Loehlin (2002) and Rowe and Rodgers (2002). *Psychological Review, 109,* 764–771.

Dolan, C. V. (1997). A note on Schönemann's refutation of Spearman's hypothesis. *Multivariate Behavioral Research, 32,* 319–325.

Dolan, C. V. (2000). Investigating Spearman's hypothesis by means of multi-group confirmatory factor analysis. *Multivariate Behavioral Research, 35,* 21–50.

Dolan, C. V., & Hamaker, E. L. (2001). Investigating Black–White differences in psychometric IQ: Multi-group confirmatory factor analyses of WISC–R and K–ABC and a critique of the method of correlated vectors. In F. Columbus (Ed.), *Advances in psychology research* (Vol. 6, pp. 30–59). Huntington, NY: Nova Science.

Eveleth, P. B., & Tanner, J. M. (1990). *Worldwide variation in human growth* (2nd ed.). Cambridge, England: Cambridge University Press.

Eyferth, K. (1961). Leistungen verscheidener Gruppen von Besatzungskindern in Ham-

burg-Wechsler Intelligenztest für Kinder (HAWIK) [Achievement of children on the Hamburg-Wechsler Intelligence Scale for Children]. *Archiv für die gesamte Psychologie, 113,* 222–241.

Eyferth, K., Brandt, U., & Hawel, W. (1960). *Farbige kinder in Deutschland* [Colored children in Germany]. Munich: Juventa Verlag.

Eysenck, H. J. (1971). *The IQ argument*. New York: Library Press.

Eysenck, H. J. (1984). The effect of race on human abilities and mental test scores. In C. R. Reynolds & R. T. Brown (Eds.), *Perspectives on bias in mental testing* (pp. 249–291). New York: Plenum.

Eysenck, H. J. (1991). Race and intelligence: An alternative hypothesis. *Mankind Quarterly, 32,* 133–136.

Fahrmeier, E. D. (1975). The effect of school attendance on intellectual development in northern Nigeria. *Child Development, 46,* 281–285.

Feuerstein, R. (1980). *The dynamic assessment of retarded performers*. Glenview, IL: Scott, Foresman.

Fisch, R. O., Bilek, M. K., Deinard, A. S., & Chang, P. N. (1976). Growth, behavioral, and psychologic measurements of adopted children: The influences of genetic and socioeconomic factors in a prospective study. *Behavioral Pediatrics, 89,* 494–500.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh, 52,* 399–433.

Flynn, J. R. (1980). *Race, IQ and Jensen*. London: Routledge & Kegan Paul.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95,* 29–51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101,* 171–191.

Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist, 54,* 5–20.

Frydman, M., & Lynn, R. (1989). The intelligence of Korean children adopted in Belgium. *Personality and Individual Differences, 10,* 1323–1326.

Galton, F. (1869). *Hereditary genius*. London: Macmillan.

Galton, F. (1883). *Inquiries into human faculty and its development*. London: Macmillan.

Gignac, G. E., Vernon, P. A., & Wickett, J. C. (2003). Factors influencing the relationship between brain size and intelligence. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 93–106). London: Elsevier.

Glewwe, P., & Jacoby, H. (1992). *Estimating the determinants of cognitive achievement in low-income countries*. Washington, DC: World Bank.

Gordon, H. L. (1934). Amentia in the East African. *Eugenics Review, 25,* 223–235.

Gordon, R. A. (1997). Everyday life as an intelligence test: Effects of intelligence and intelligence context. *Intelligence, 24,* 203–320.

Gottfredson, L. S. (1997). Why *g* matters: The complexity of everyday life. *Intelligence, 24,* 79–132.

Gottfredson, L. S. (2000). Pretending that intelligence doesn't matter. *Cerebrum, 2,* 75–96.

Gould, S. J. (1996). *The mismeasure of man* (2nd ed.). New York: Norton.

Grieve, K. W., & Viljoen, S. (2000). An exploratory study of the use of the Austin Maze in South Africa. *South African Journal of Psychology, 30,* 14–18.

Haier, R. J., Chueh, D., Touchette, P., Lott, I., Buchsbaum, M., Macmillan, D., et al. (1995). Brain size and cerebral glucose metabolic rate in nonspecific mental retardation and Down syndrome. *Intelligence, 20,* 191–210.

Harvey, I., Persaud, R., Ron, M. A., Baker, G., & Murray, R. M. (1994). Volumetric MRI measurements in bipolars compared with schizophrenics and healthy controls. *Psychological Medicine, 24,* 689–699.

Herrnstein, R. J. (1990). Still an American dilemma. *Public Interest, 98,* 3–17.
Herrnstein, R. J., & Murray, C. (1994). *The bell curve*. New York: Free Press.
Hills, J. R., & Stanley, J. C. (1970). Easier test improves prediction of Black students' college grades. *Journal of Negro Education, 39,* 320–324.
Ho, K. C., Roessmann, U., Straumfjord, J. V., & Monroe, G. (1980). Analysis of brain weight: I and II. *Archives of Pathology and Laboratory Medicine, 104,* 635–645.
Hoberman, J. M. (1997). *Darwin's athletes: How sport has damaged Black America and preserved the myth of race*. New York: Houghton Mifflin.
Howells, W. W. (1993). *Getting here: The story of human evolution*. Washington, DC: Compass Press.
Jencks, C., & Phillips, M. (Eds.). (1998). *The Black–White test score gap*. Washington, DC: Brookings Institution Press.
Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39,* 1–123.
Jensen, A. R. (1973). *Educability and group differences*. London: Methuen.
Jensen, A. R. (1974). Cumulative deficit: A testable hypothesis? *Developmental Psychology, 10,* 996–1019.
Jensen, A. R. (1977). Cumulative deficit in IQ of Blacks in the rural South. *Developmental Psychology, 13,* 184–191.
Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.
Jensen, A. R. (1992). Spearman's hypothesis: Methodology and evidence. *Multivariate Behavioral Research, 27,* 225–234.
Jensen, A. R. (1993). Spearman's hypothesis tested with chronometric information processing tasks. *Intelligence, 17,* 47–77.
Jensen, A. R. (1994). Psychometric *g* related to differences in head size. *Personality and Individual Differences, 17,* 597–606.
Jensen, A. R. (1997). The puzzle of nongenetic variance. In R. J. Sternberg & E. L. Grigorenko (Eds.), *Intelligence, heredity, and environment* (pp. 42–88). Cambridge, England: Cambridge University Press.
Jensen, A. R. (1998a). Adoption data and two *g*-related hypotheses. *Intelligence, 25,* 1–6.
Jensen, A. R. (1998b). *The g factor*. Westport, CT: Praeger.
Jensen, A. R. (2000). Testing: The dilemma of group differences. *Psychology, Public Policy, and Law, 6,* 121–127.
Jensen, A. R., & Johnson, F. W. (1994). Race and sex differences in head size and IQ. *Intelligence, 18,* 309–333.
Jensen, A. R., & Reynolds, C. R. (1982). Race, social class and ability patterns on the WISC–R. *Personality and Individual Differences, 3,* 423–438.
Jensen, A. R., & Whang, P. A. (1994). Speed of accessing arithmetic facts in long-term memory: A comparison of Chinese-American and Anglo-American children. *Contemporary Educational Psychology, 19,* 1–12.
Kamin, L. (1974). *The science and politics of IQ*. Hillsdale, NJ: Erlbaum.
Kendall, I. M., Verster, M. A., & von Mollendorf, J. W. (1988). Test performance of Blacks in Southern Africa. In S. H. Irvine & J. W. Berry (Eds.), *Human abilities in cultural context* (pp. 299–339). Cambridge, England: Cambridge University Press.
Klitgaard, R. (1986). *Elitism and meritocracy in developing countries: Selection policies for higher education.* Baltimore: Johns Hopkins University Press.
Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–195). Cambridge, England: Cambridge University Press.
Lakatos, I. (1978). *The methodology of scientific research programs*. Cambridge, England: Cambridge University Press.

Levin, M. (1994). Comment on Minnesota Transracial Adoption Study. *Intelligence, 19,* 13–20.

Levin, M. (1997). *Why race matters*. Westport, CT: Praeger.

Lewontin, R. C., Rose, S., & Kamin, L. J. (1984). *Not in our genes*. New York: Pantheon.

Loehlin, J. C. (2000). Group differences in intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 176–193). Cambridge, England: Cambridge University Press.

Loehlin, J. C. (2002). The IQ paradox: Resolved? Still an open question. *Psychological Review, 109,* 754–758.

Loehlin, J. C., Lindzey, G., & Spuhler, J. N. (1975). *Race differences in intelligence*. San Francisco: Freeman.

Loehlin, J. C., Vandenberg, S. G., & Osborne, R. T. (1973). Blood group genes and Negro–White differences. *Behavioral Genetics, 3,* 263–270.

Loury, G. C. (2002). *The anatomy of racial inequality*. Cambridge, MA: Harvard University Press.

Lynn, R. (1977). The intelligence of the Japanese. *Bulletin of the British Psychological Society, 30,* 69–72.

Lynn, R. (1978). Ethnic and racial differences in intelligence: International comparisons. In R. T. Osborne, C. E. Noble, & N. Weyl (Eds.), *Human variation: The biopsychology of age, race, and sex* (pp. 261–286). New York: Academic Press.

Lynn, R. (1982, May 20). IQ in Japan and the United States shows a growing disparity. *Nature, 297,* 222–223.

Lynn, R. (1990). The role of nutrition in secular increase in intelligence. *Personality and Individual Differences, 11,* 273–285.

Lynn, R. (1991). Race differences in intelligence: A global perspective. *Mankind Quarterly, 31,* 255–296.

Lynn, R. (1994). Some reinterpretations of the Minnesota Transracial Adoption Study. *Intelligence, 19,* 21–27.

Lynn, R. (1996). Racial and ethnic differences in intelligence in the United States on the Differential Ability Scale. *Personality and Individual Differences, 20,* 271–273.

Lynn, R. (2002). Skin color and intelligence in African Americans. *Population and Environment, 23,* 365–375.

Lynn, R., & Hattori, K. (1990). The heritability of intelligence in Japan. *Behavior Genetics, 20,* 545–546.

Lynn, R., & Holmshaw, M. (1990). Black–White differences in reaction times and intelligence. *Social Behavior and Personality, 18,* 299–308.

Lynn, R., & Owen, K. (1994). Spearman's hypothesis and test score differences between Whites, Indians, and Blacks in South Africa. *Journal of General Psychology, 121,* 27–36.

Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations.* Westport, CT: Praeger.

MacArthur, R. S. (1968). Some differential abilities of northern Canadian native youth. *International Journal of Psychology, 3,* 43–51.

Mall, F. P. (1909). On several anatomical characters of the human brain, said to vary according to race and sex, with special reference to the weight of the frontal lobe. *American Journal of Anatomy, 9,* 1–32.

Masters, R. (1997). Brain biochemistry and social status: The neurotoxity hypothesis. In E. White (Ed.), *Intelligence, political inequality, and public policy* (pp. 141–183). Westport, CT: Praeger.

McGue, M., Bouchard, T. J., Jr., Iacono, W. G., & Lykken, D. T. (1993). Behavioral genetics of cognitive ability: A life-span perspective. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture, and psychology* (pp. 59–76). Washington, DC: American Psychological Association.

Mealey, L. (1990). Differential use of reproductive strategies by human groups? *Psychological Science, 1,* 385–387.

Miele, F. (1995). For whom the bell curve tolls: An interview with [the author of *The Bell Curve*] Charles Murray, on IQ, race, class, Gould, Gardner, and the Clintons. *Skeptic, 3*(2), 34–41.

Miele, F. (2002). *Intelligence, race, and genetics: Conversations with Arthur R. Jensen.* Boulder, CO: Westview Press.

Miller, E. M. (1994). Intelligence and brain myelination: A hypothesis. *Personality and Individual Differences, 17,* 803–832.

Moore, E. G. J. (1986). Family socialization and the IQ test performance of traditionally and transracially adopted Black children. *Developmental Psychology, 22,* 317–326.

Morton, S. G. (1849). Observations on the size of the brain in various races and families of man. *Proceedings of the Academy of Natural Sciences, Philadelphia, 4,* 221–224.

Myrdal, G. (1944). *An American dilemma*. New York: Harper & Row.

Naglieri, J. A., & Jensen, A. R. (1987). Comparison of Black–White differences on the WISC–R and K–ABC: Spearman's hypothesis. *Intelligence, 11,* 21–43.

Nagoshi, C. T., Johnson, R. C., DeFries, J. C., Wilson, J. R., & Vandenberg, S. G. (1984). Group differences and first principal-component loadings in the Hawaii Family Study of Cognition: A test of the generality of "Spearman's hypothesis." *Personality and Individual Differences, 5,* 751–753.

Nei, M., & Roychoudhury, A. K. (1993). Evolutionary relationships of human populations on a global scale. *Molecular Biology and Evolution, 10,* 927–943.

Neisser, U. (1997). Never a dull moment. *American Psychologist, 52,* 79–81.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51,* 77–101.

Nell, V. (2000). *Cross-cultural neuropsychological assessment: Theory and practice.* Mahwah, NJ: Erlbaum.

Nisbett, R. E. (1998). Race, genetics, and IQ. In C. Jencks & M. Phillips (Eds.), *The Black–White test score gap* (pp. 86–102). Washington, DC: Brookings Institute.

Nott, J. C., & Glidden, G. R. (1854). *Types of mankind*. Philadelphia: Lippincott.

Nyborg, H. (1994). *Hormones, sex, and society: The science of physicology*. Westport, CT: Praeger.

Nyborg, H., & Jensen, A. R. (2000). Black–White differences on various psychometric tests: Spearman's hypothesis tested on American Armed Services veterans. *Personality and Individual Differences, 28,* 593–599.

Ogbu, J. U. (2002). Cultural amplifiers of intelligence: IQ and minority status in cross-cultural perspective. In J. M. Fish (Ed.), *Race and intelligence: Separating science from myth* (pp. 241–278). Mahwah, NJ: Erlbaum.

Osborne, R. T. (1980). *Twins: Black and White*. Athens, GA: Foundation for Human Understanding.

Osborne, R. T., & McGurk, F. C. J. (Eds.). (1982). *The testing of Negro intelligence: Vol. 2*. Athens, GA: Foundation for Human Understanding.

Owen, K. (1992). The suitability of Raven's Standard Progressive Matrices for various groups in South Africa. *Personality and Individual Differences, 13,* 149–159.

Papiernik, E., Cohen, H., Richard, A., de Oca, M. M., & Feingold, J. (1986). Ethnic differences in duration of pregnancy. *Annals of Human Biology, 13,* 259–265.

Parra, E. J., Marcini, A., Akey, J., Martinson, J., Batzer, M. A., Cooper, R., et al. (1998). Estimating African American admixture proportions by use of population specific alleles. *American Journal of Human Genetics, 63,* 1839–1851.

Pearl, R. (1934, May 11). The weight of the Negro brain. *Science, 80,* 431–434.

Pennington, B. F., Filipek, P. A., Lefly, D., Chhabildas, N., Kennedy, D. N., Simon, J. H.,

et al. (2000). A twin MRI study of size variations in the human brain. *Journal of Cognitive Neuroscience, 12,* 223–232.

Peoples, C. E., Fagan, J. F., III, & Drotar, D. (1995). The influence of race on 3-year-old children's performance on the Stanford–Binet: Fourth edition. *Intelligence, 21,* 69–82.

Platt, J. R. (1964, January 24). Strong inference. *Science, 146,* 347–353.

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences, 10,* 1–60.

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype–environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84,* 309–322.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2001). *Behavioral genetics* (4th ed.). New York: Freeman.

Polednak, A. P. (1989). *Racial and ethnic differences in disease*. Oxford, England: Oxford University Press.

Post, R. H. (1982). Populational differences in visual acuity: A review, with speculative notes on selection relaxation. *Social Biology, 29,* 319–343.

Posthuma, D., De Geus, E. J. C., Baare, W. F. C., Pol, H. E. H., Kahn, R. S., & Boomsma, D. I. (2002). The association between brain volume and intelligence is of genetic origin. *Nature Neuroscience, 5,* 83–84.

Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology, 41,* 1–48.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Standard Progressive Matrices* (1998 ed.). Oxford, England: Oxford Psychologists Press.

Raven, J., Summers, B., Birchfield, M., Brosier, G., Burciaga, L., Byrkit, B., et al. (1990). *Manual for Raven's Progressive Matrices and Vocabulary Scales—Research supplement: No. 3. American and international norms* (2nd ed.). Oxford, England: Oxford Psychologists Press.

Redding, R. E. (2001). Sociopolitical diversity in psychology: The case for pluralism. *American Psychologist, 56,* 205–215.

Ree, M. J., & Carretta, T. R. (1995). Group differences in aptitude factor structure on the ASVAB. *Educational and Psychological Measurement, 55,* 268–277.

Risch, N., Burchard, E., Ziv, E., & Tang, H. (2002). Categorization of humans in biomedical research: Genes, race, and disease. *Genome Biology, 3*(7), 1–12.

Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., III, & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology, 54,* 297–330.

Rowe, D. C. (1994). No more than skin deep. *American Psychologist, 49,* 215–216.

Rowe, D. C. (2002). IQ, birth weight, and number of sexual partners in White, African American, and mixed race adolescents. *Population and Environment, 23,* 513–524.

Rowe, D. C., & Cleveland, H. H. (1996). Academic achievement in Blacks and Whites: Are the developmental processes similar? *Intelligence, 23,* 205–228.

Rowe, D. C., & Rodgers, J. L. (2002). Expanding variance and the case of historical changes in IQ means: A critique of Dickens and Flynn (2001). *Psychological Review, 109,* 759–763.

Rowe, D. C., Vazsonyi, A. T., & Flannery, D. J. (1994). No more than skin deep: Ethnic and racial similarity in developmental process. *Psychological Review, 101,* 396–413.

Rowe, D. C., Vazsonyi, A. T., & Flannery, D. J. (1995). Ethnic and racial similarity in developmental process: A study of academic achievement. *Psychological Science, 6,* 33–38.

Rushton, J. P. (1989). Japanese inbreeding depression scores: Predictors of cognitive differences between Blacks and Whites. *Intelligence, 13,* 43–51.

Rushton, J. P. (1991). Mongoloid–Caucasoid differences in brain size from military samples. *Intelligence, 15,* 351–359.

Rushton, J. P. (1992). Cranial capacity related to sex, rank, and race in a stratified random sample of 6,325 U.S. military personnel. *Intelligence, 16,* 401–413.

Rushton, J. P. (1993). Corrections to a paper on race and sex differences in brain size and intelligence. *Personality and Individual Differences, 15,* 229–231.

Rushton, J. P. (1994). Sex and race differences in cranial capacity from International Labour Office data. *Intelligence, 19,* 281–294.

Rushton, J. P. (1997). Cranial size and IQ in Asian Americans from birth to age seven. *Intelligence, 25,* 7–20.

Rushton, J. P. (1999). Secular gains in IQ not related to the *g* factor and inbreeding depression—unlike Black–White differences: A reply to Flynn. *Personality and Individual Differences, 26,* 381–389.

Rushton, J. P. (2000). *Race, evolution, and behavior: A life history perspective* (3rd ed.). Port Huron, MI: Charles Darwin Research Institute.

Rushton, J. P. (2001). Black–White differences on the *g* factor in South Africa: A "Jensen effect" on the Wechsler Intelligence Scale for Children—Revised. *Personality and Individual Differences, 31,* 1227–1232.

Rushton, J. P. (2002). Jensen effects and African/Coloured/Indian/White differences on Raven's Standard Progressive Matrices in South Africa. *Personality and Individual Differences, 33,* 1279–1284.

Rushton, J. P., & Ankney, C. D. (1996). Brain size and cognitive ability: Correlations with age, sex, social class, and race. *Psychonomic Bulletin and Review, 3,* 21–36.

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin, 94,* 18–38.

Rushton, J. P., & Jensen, A. R. (2003). African–White IQ differences from Zimbabwe on the Wechsler Intelligence Scale for Children—Revised are mainly on the *g* factor. *Personality and Individual Differences, 34,* 177–183.

Rushton, J. P., & Osborne, R. T. (1995). Genetic and environmental contributions to cranial capacity in Black and White adolescents. *Intelligence, 20,* 1–13.

Rushton, J. P., & Skuy, M. (2000). Performance on Raven's Matrices by African and White university students in South Africa. *Intelligence, 28,* 251–265.

Rushton, J. P., Skuy, M., & Fridjhon, P. (2002). Jensen effects among African, Indian, and White engineering students in South Africa on Raven's Standard Progressive Matrices. *Intelligence, 30,* 409–423.

Rushton, J. P., Skuy, M., & Fridjhon, P. (2003). Performance on Raven's Advanced Progressive Matrices by African, East Indian, and White engineering students in South Africa. *Intelligence, 31,* 123–137.

Scarr, S. (1996). How people make their own environments: Implications for parents and policy makers. *Psychology, Public Policy, and Law, 2,* 204–228.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype–environment effects. *Child Development, 54,* 424–435.

Scarr, S., Pakstis, A. J., Katz, S. H., & Barker, W. B. (1977). Absence of a relationship between degree of White ancestry and intellectual skills within a Black population. *Human Genetics, 39,* 69–86.

Scarr, S., & Weinberg, R. A. (1976). IQ test performance of Black children adopted by White families. *American Psychologist, 31,* 726–739.

Scarr-Salapatek, S. (1971, December 24). Race, social class, and IQ. *Science, 174,* 1285–1295.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262–274.

Schönemann, P. H. (1992). Extension of Guttman's result from *g* to PC1. *Multivariate Behavioral Research, 27,* 219–224.

Schoenemann, P. T., Budinger, T. F., Sarich, V. M., & Wang, W. S. Y. (2000). Brain size does not predict general cognitive ability within families. *Proceedings of the National Academy of Sciences USA, 97,* 4932–4937.

Shigehisa, T., & Lynn, R. (1991). Reaction times and intelligence in Japanese children. *International Journal of Psychology, 26,* 195–202.

Shuey, A. M. (1958). *The testing of Negro intelligence*. Lynchburg, VA: J. P. Bell.

Shuey, A. M. (1966). *The testing of Negro intelligence* (2nd ed.). New York: Social Science Press.

Simmons, K. (1942). Cranial capacities by both plastic and water techniques with cranial linear measurements of the Reserve Collection: White and Negro. *Human Biology, 14,* 473–498.

Skuy, M., Gewer, A., Osrin, Y., Khunou, D., Fridjhon, P., & Rushton, J. P. (2002). Effects of mediated learning experience on Raven's Matrices scores of African and non-African university students in South Africa. *Intelligence, 30,* 221–232.

Skuy, M., Hoffenberg, S., Visser, L., & Fridjhon, P. (1990). Temperament and cognitive modifiability of academically superior Black adolescents in South Africa. *International Journal of Disability, Development, and Education, 37,* 29–44.

Skuy, M., Schutte, E., Fridjhon, P., & O'Carroll, S. (2001). Suitability of published neuropsychological test norms for urban African secondary school students in South Africa. *Personality and Individual Differences, 30,* 1413–1425.

Skuy, M., & Shmukler, D. (1987). Effectiveness of the Learning Potential Assessment Device with Indian and Colored adolescents in South Africa. *International Journal of Special Education, 12,* 131–149.

Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist*, *42,* 137–144.

Snyderman, M., & Rothman, S. (1988). *The IQ controversy, the media, and public policy*. New Brunswick, NJ: Transaction Books.

Sowell, T. (1994). *Race and culture*. New York: Basic Books.

Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.

Stanley, J. C., & Porter, A. C. (1967). Correlation of Scholastic Aptitude Test scores with college grades for Negroes versus Whites. *Journal of Educational Measurement, 4,* 199–218.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52,* 613–629.

Sternberg, R. J. (2000). (Ed.). *Handbook of intelligence*. Cambridge, England: Cambridge University Press.

Sternberg, R. J., Nokes, C., Geissler, P. W., Prince, R., Okatcha, F., Bundy, D. A., & Grigorenko, E. L. (2001). The relationship between academic and practical intelligence: A case study in Kenya. *Intelligence, 29,* 401–418.

Stringer, C., & McKie, R. (1996). *African exodus*. London: Cape.

Tate, C., & Audette, D. (2001). Theory and research on "race" as a natural kind variable in psychology. *Theory and Psychology, 11,* 495–520.

te Nijenhuis, J., Voskuijl, O. F., & Schijve, N. B. (2001). Practice and coaching on IQ tests: Quite a lot of *g*. *International Journal of Selection and Assessment, 9,* 302–308.

Thompson, P. M., Cannon, T. D., Narr, K. L., Erp, T. V., Poutanen, V. P. Huttunen, M., et al. (2001). Genetic influences on brain structure. *Nature Neuroscience, 4,* 1253–1258.

Tizard, B. (1974, February 1). IQ and race. *Nature, 247,* 316.

Todd, T. W. (1923). Cranial capacity and linear dimensions, in White and Negro. *American Journal of Physical Anthropology, 6,* 97–194.

Urbach, P. (1974a). Progress and degeneration in the "IQ debate" (I). *British Journal of the Philosophy of Science, 25,* 99–135.

Urbach, P. (1974b). Progress and degeneration in the "IQ debate" (II). *British Journal of the Philosophy of Science, 25,* 235–259.

Vernon, P. A., Wickett, J. A., Bazana, G., & Stelmack, R. M. (2000). The neuropsychology and psychophysiology of human intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 245–264). Cambridge, England: Cambridge University Press.

Vernon, P. E. (1965). Ability factors and environmental influences. *American Psychologist, 20,* 723–733.

Vernon, P. E. (1979). *Intelligence: Heredity and environment*. San Francisco: Freeman.

Vernon, P. E. (1982). *The abilities and achievements of Orientals in North America*. New York: Academic Press.

Vint, F. W. (1934). The brain of the Kenya native. *Journal of Anatomy, 48,* 216–223.

Waldman, I. D., Weinberg, R. A., & Scarr, S. (1994). Racial-group differences in IQ in the Minnesota Transracial Adoption Study: A reply to Levin and Lynn. *Intelligence, 19,* 29–44.

Watson, J. B. (1924). *Behaviorism*. Chicago: University of Chicago Press.

Weinberg, R. A., Scarr, S., & Waldman, I. D. (1992). The Minnesota Transracial Adoption Study: A follow-up of IQ test performance at adolescence. *Intelligence, 16,* 117–135.

Wigdor, A. K., & Garner, W. R. (Eds.). (1982). *Ability testing: Uses, consequences, and controversies. Part 1: Report of the Committee. Part 2: Documentation section.* Washington, DC: National Academy Press.

Willerman, L., Naylor, A. F., & Myrianthopoulos, N. C. (1974). Intellectual development of children from interracial matings: Performance in infancy and at 4 years. *Behavioral Genetics, 4,* 84–88.

Wilson, E. O. (1978). *On human nature*. Cambridge, MA: Harvard University Press.

Winick, M., Meyer, K. K., & Harris, R. C. (1975, December 19). Malnutrition and environmental enrichment by early adoption. *Science, 190,* 1173–1175.

Zaaiman, H., van der Flier, H., & Thijs, G. D. (2001). Dynamic testing in selection for an educational programme: Assessing South African performance on the Raven Progressive Matrices. *International Journal of Selection and Assessment, 9,* 258–269.

Zindi, F. (1994). Differences in performance. *The Psychologist, 7,* 549–552.

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at http://watson.apa.org/notify/ and you will be notified by e-mail when issues of interest to you become available!